



# International Conference on Life Sciences, Engineering and Technology

09 - 12 April 2026  
New York, USA



# Proceedings of International Conference on Life Sciences, Engineering and Technology

Volume: 7

© 2026 Published by ISTES

ISSN: 2833-0447

The proceedings is licensed under a Creative Commons Attribution-NonCommercialShareAlike 4.0 International License, permitting all non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited. Authors alone are responsible for the contents of their papers. The Publisher, the ISTES, shall not be liable for any loss, actions, claims, proceedings, demand, or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of the research material. All authors are requested to disclose any actual or potential conflict of interest including any financial, personal or other relationships with other people or organizations regarding the submitted work. The submissions are subject to a double-blind peer review process by at least two reviewers with expertise in the relevant subject area. The review policy is available on the conference web page.

Conference: International Conference on Life Sciences, Engineering and Technology (ILSET)

Conference Dates: 09 - 12 April 2026

Conference Location: New York, USA

Publication Date: December 30, 2026

## Editors

Dr. Mack SHELLEY - Iowa State University, United States

Dr. Wilfried ADMIRAAL - Leiden University, the Netherlands

Dr. I-Tsun CHIANG - National Taiwan Normal University, Taiwan

Dr. Wenxia WU - George Washington University, United States

## CONFERENCE PRESIDENTS

Dr. Mack SHELLEY - Iowa State University, United States  
Dr. Wilfried ADMIRAAL - Leiden University, the Netherlands  
Dr. I-Tsun CHIANG - National Taiwan Normal University, Taiwan  
Dr. Wenxia WU - George Washington University, United States

## SCIENTIFIC BOARD

Allan TARP - MATHeCADEMY, Denmark  
Altay FIRAT - European University of Lefke, Cyprus  
Andrea DEBELJUH - University Juraj Dobrila of Pula, Croatia  
Brahim FERDI - Bechar University, Algeria  
Branislav POPOVIĆ - University of Kragujevac, Serbia  
Chalavadi SULOCHANA - Gulbarga University, India  
Dariga NURKESHEVA - Nazarbayev University, Kazakhstan  
Elizabeth ADAMSON - Edinburgh Napier University, United Kingdom  
Farhad BALASH - Kharazmi University, Iran  
Farouq ALMEQDADI - Emirates College for Advanced Education (ECAE), U.A.E.  
Gordana SAVIC - University of Belgrade, Serbia  
Haydar YUKSEK - Kafkas University, Turkey  
Henry David KATNIYON - Federal College of Education, Pankshin, Plateau state, Nigeria  
Hsin-Chih WU - National Taiwan Normal University, Taiwan  
Janet VALDEZ - La Consolacion University Philippines, Philippines  
Jessie BUSTILLOS - London Metropolitan University, United Kingdom  
Kartika YULIANTI - Bina Nusantara University, Indonesia  
Mesut AYDIN - Inonu University, Turkey  
Milica PAVKOV HRVOJEVIĆ - University of Novi Sad, Serbia  
Mohammad SARWAR - Scialert, Dubai, United Arab Emirates  
Morteza BARIN - Farhangiyen University of Iran, Iran  
Muteb ALQAHTANI - Rutgers University, United States  
Nurten SARGIN - Necmettin Erbakan University, Turkey  
Ognyan B. MANOLOV - European Polytechnic University, Bulgaria  
Sanaa AL-DELAIFY - Mosul University, Iraq  
Shynar BAIMAGANBETOVA - Nazarbayev University, Kazakhstan  
Summer MOUALLEM - University of Central Lancashire, United Kingdom  
Tamer USTUNER - Kahramanmaraş Sutcu Imam University, Turkey  
Tri Marhaeni PUDJI ASTUTI - Semarang State University, Indonesia  
Yi-Teng HSU - National Taiwan Normal University, Taiwan

## ORGANIZING COMMITTEE

Aliya MUSTAFINA - Nazarbayev University, Kazakhstan  
Ann D. THOMPSON - Iowa State University, United States  
Halil SNOPCE - South East European University, Macedonia  
Hakan AKCAY - Yildiz Technical University, Turkey  
Jacqueline T. MCDONNOUGH - Virginia Commonwealth University, United States  
Mariusz JAWORSKI - Medical University of Warsaw, Poland  
Mary M. CAPRARO - Texas A&M University, United States  
Muhammad ZAYYAD - Al-Qasemi Academic College of Education, Israel  
Murat BEYTUR - Kafkas University, Turkey  
Mustafa PEHLIVAN - Necmettin Erbakan University, Turkey  
Natela DOGHONADZE - International Black Sea University, Georgia  
Natalija ACESKA - Ministry of Education and Science, Macedonia  
Ossi AUTIO - University of Helsinki, Finland  
O. Tayfur OZTURK - Necmettin Erbakan University, Turkey  
Silvia MORARU - National High School Bucharest, Romania  
Silviya KOSTOVA - The D.A. Tsenov Academy of Economics, Bulgaria

## PARTICIPATING COUNTRIES

|                                |             |         |
|--------------------------------|-------------|---------|
| United States of America (USA) | South Korea | Poland  |
| Kuwait                         | CS-USA      | Bahrain |
| Israel                         | Puerto Rico | Albania |
| Italy                          | Lithuania   | Nigeria |
| Brazil                         | Türkiye     |         |

## INDEXES

The publications affiliated with ISTES Organization are indexed or listed by all or some of the following sources:



## SUPPORTED BY



## ISTES EVENTS



### International Conference on Studies in Education and Social Sciences

05 - 09 November 2026 - Istanbul, Türkiye



### International Conference on Studies in Engineering, Science, and Technology

05 - 09 November 2026 - Istanbul, Türkiye



### International Conference on Social, Educational, and Human sciences

04 - 07 December 2026 - Istanbul, Türkiye



### International Conference on Engineering, Technology, and Life sciences

04 - 07 December 2026 - Istanbul, Türkiye



### International Conference on Behavioral, Educational, and Social Sciences

04 - 07 December 2026 - Miami, USA



### International Conference on Business, Engineering, Science, and Technology

04 - 07 December 2026 - Miami, USA



### International Conference on Business, Art, Social sciences, and Education

04 - 07 December 2026 - Paris, France



### International Conference on Technology, Applied Sciences, and Engineering

04 - 07 December 2026 - Paris, France



### International Conference on Humanities, Social and Education Sciences

08 - 11 April 2027 - Los Angeles, USA



### International Conference on Life Sciences, Engineering and Technology

08 - 11 April 2027 - Los Angeles, USA



### International Conference on Research in Education and Science

28 April - 02 May 2027 - Marmaris, Muğla / Türkiye

## ISTES EVENTS



### International Conference on Education in Mathematics, Science and Technology

28 April - 02 May 2027 - Marmaris, Muğla / Türkiye



### International Conference on Humanities, Education, and Social Sciences

16 - 19 March 2027 - Singapore



### International Conference on Technology, Engineering, and Life Sciences

16 - 19 March 2027 - Singapore



### International Conference on Social science, Education science, and Humanities

26 - 29 July 2027 - Wahsington, USA



### International Conference on Engineering, Life sciences, and Technology

26 - 29 July 2027 - Wahsington, USA



### International Conference on Humanities and Social Sciences

26 - 29 July 2027 - Vienna, Austria



### International Conference on Life Sciences and Engineering

26 - 29 July 2027 - Vienna, Austria



### International Conference on Humanities, Education, and Social sciences

26 - 29 July 2027 - Singapore



### International Conference on Health, Engineering, and Life sciences

26 - 29 July 2027 - Singapore



### International Conference on Engineering, Science and Technology

21 - 24 October 2027 - Washington D.C., USA



### International Conference on Social and Education Sciences

21 - 24 October 2027 - Washington D.C., USA

## Contents

|                                                                                                                                           |    |
|-------------------------------------------------------------------------------------------------------------------------------------------|----|
| Online Assessment Integrity in the Era of Generative AI: A Comparative Study of Student and AI-Generated Responses in Online Quizzes..... | 1  |
| GenAI-Assisted Cybersecurity Mesh for Secure Digital Transactions.....                                                                    | 12 |
| A Qualitative Study of University Student Experiences Regarding the Impact of Artificial Intelligence on Coding Coursework.....           | 25 |
| A Constitution for Synthetic Artificial Intelligence Systems: A Wisdom-Based and Sociotechnical Approach..                                | 31 |
| An Examination of Fine Arts Faculty Students' Perceptions of Pedagogical Art Criticism Competence.....                                    | 37 |
| Branding Without Belonging: How AI-Generated Content Marketing Erases the Cultural Context of Heritage Foods.....                         | 48 |



## Online Assessment Integrity in the Era of Generative AI: A Comparative Study of Student and AI-Generated Responses in Online Quizzes

Yavuz Keceli <sup>1\*</sup>, Andrew W. Maguire <sup>2</sup>, Jack C. Meyer <sup>3</sup>

<sup>1</sup> Alfred University, USA, , 0000-0002-8981-8217

<sup>2</sup> Alfred University, USA, 0009-0000-1260-2007

<sup>3</sup> Alfred University, USA, 0009-0009-7195-8340

\* Corresponding author: Yavuz Keceli (ilsetoffice@gmail.com)

### Article Info

#### Article History

Received:

24 February 2026

Revised:

25 April 2026

Accepted:

11 July 2026

Published:

30 August 2026

#### Keywords

Artificial intelligence

Higher education

Online quizzes

Assessment reliability

### Abstract

In the rapidly evolving era of generative artificial intelligence (GenAI), higher education faces an unprecedented challenge regarding the integrity and validity of traditional assessment methods. This is particularly evident in online environments where unsupervised quizzes allow students easy access to GenAI tools. Therefore, the purpose of this study is to test the reliability of online quizzes and determine which question types best reflect a student's true abilities without the influence of AI assistance. In a controlled experimental environment, student volunteers participated in a mock quiz featuring ten unique multiple-choice questions, answering first independently and subsequently with AI assistance. Some of these questions were selected from a textbook publisher's test bank, whereas some questions were specifically crafted to refer to the classroom discussions that were not part of any textbook. Student perceptions and experiences were further evaluated through pre-test and post-test feedback surveys. Based on the results, this paper proposes several recommendations for educators to improve the quality and security of assessments in online education.

**Citation:** Keceli, Y., Maguire, A. W., & Meyer, J. C. (2026). Online Assessment Integrity in the Era of Generative AI: A Comparative Study of Student and AI-Generated Responses in Online Quizzes. In M. Shelley, S. Turgut, & E. Cakir (Eds.), *Proceedings of International Conference on Life Sciences, Engineering and Technology* (pp. 1–11), New York, USA. ISTES.



Volume: 7, © 2026, Published by ISTES, ISBN: 978-1-952092-79-4, ISSN: 2833-0447



## Introduction

The rapid shift toward digital learning environments has transformed how higher education institutions evaluate student performance. While online assessments offer flexibility and efficiency, they have also introduced significant challenges regarding academic honesty. Research suggests that the digital environment itself may influence student behavior; for instance, students are significantly more likely to engage in misconduct, such as obtaining answers from peers, during online assessments compared to traditional in-person examinations (Holden et al., 2021).

The emergence of generative artificial intelligence has further complicated this landscape. Students may now exploit these tools during online exams to retrieve answers in real-time, effectively bypassing the intended evaluation process. This shift toward AI-assisted misconduct poses a serious threat to the fairness, reliability, and overall accuracy of academic assessments (Gundu, 2025).

This paper explores these challenges through a controlled experiment. By analyzing how freshmen business students interact with AI during a quiz, this study examines whether specific question structures can mitigate the influence of generative AI and Large Language Models (LLMs). Through a combination of statistical analysis and qualitative student feedback, we aim to identify practical strategies for designing "AI-resilient" assessments that encourage authentic learning and student participation.

## Theoretical Background

### Test and Exam Design

Bennet (2026) examines Massive Open Online Courses (MOOCs) and argues that detection-first approaches are unfeasible at scale; not only is it impossible to reliably identify all instances of misconduct, but false positives also risk harming innocent students. Consequently, the focus must shift toward AI-resilient online assessment design, key principles of which include assigning authentic, real-world tasks, requiring documentation of the creative process, and tailoring questions to individual student experiences. Furthermore, incorporating oral components, such as video presentations or live discussions, and utilizing iterative submissions with multiple feedback loops can further safeguard integrity.

This perspective is supported by Bittle and El-Gayar (2025), who suggest that redesigning assessments to promote higher-order thinking is crucial. They argue that traditional methods are often ineffective in a generative AI-enhanced environment and advocate for project-based assessments and oral exams that challenge students to demonstrate creativity and problem-solving. Similarly, in an investigation into engineering education, Nikolic et al. (2023) note that while tools like ChatGPT may not answer every quiz question correctly, they often provide enough accuracy for a student to pass. To mitigate this, they propose designing questions that incorporate images and tables to make them harder to simply copy and paste, and focusing on specific content from lecture notes rather than generic concepts.

Evangelista (2025) provides a systematic review of best practices, highlighting strategies such as interactive feedback mechanisms, the use of multimedia in prompts, and assessing the learning process rather than merely the final product. This shift toward process-based evaluation is echoed by Hanh and Duyen (2025); their qualitative study of postgraduate students revealed that reinforcing exam design and supervision, alongside establishing ethical frameworks and "communities of integrity," are perceived as primary preventative measures.

While some focus on design, others evaluate technical and structural safeguards. Holden et al. (2021) note that advanced detection tools, such as biometrics, often raise significant privacy concerns and may be unfeasible. Instead, they recommend diversifying the structure and delivery of exam items.

Kara-Mohamad (2025) evaluates specific techniques for Virtual Learning Environments (VLEs) like Canvas, such as utilizing large question banks, replacing text with images, and using unpublished materials. While these methods minimize risk, Kara-Mohamad acknowledges they cannot provide a total guarantee of integrity.

Cantiello and Geschke (2024) offer complementary practical recommendations, including the use of time limits, question randomization, lockdown browsers, and treating online quizzes as low-stakes assignments.

Finally, the shift away from traditional testing is supported by instructional designers. Martin, F., Kim, S., Bolliger, D. U., et al (2025) found that these professionals rank both proctored and non-proctored online exams as the least effective assessment methods compared to case studies, electronic portfolios, and design projects.

## **Pandemic**

The COVID-19 pandemic forced a lot of academic institutions into online education for an extended period of time, sparking significant concern regarding the quality and integrity of remote assessments. Uca Uça Güneş et al. (2024) surveyed 163 instructors who taught during this period, finding that integrity measures were largely dictated by the technical capabilities of the institution's Learning Management System (LMS). Their findings highlight that the potential for ethical violations remains a significant threat, necessitating the urgent identification and integration of robust solutions into online examination workflows.

Categorizing the broader landscape of academic dishonesty, Noorbehbahani et al. (2022) conducted a systematic review of 58 publications, organizing the literature into four key domains: reasons for cheating, types of misconduct, detection methods, and prevention strategies. A primary prevention method identified is the design of "cheat-resistant" exams. This approach utilizes personalized, knowledge-based questions—which require higher-level synthesis—rather than simple information-based questions that are easily searched (p. 8439).

The effectiveness of delivery formats also remains a point of debate. Hartley et al. (2025) compared student performance across three settings: in-person, unmonitored online, and online with video monitoring. While they noted significant differences in grades across these formats, the authors argue that performance variance alone

cannot be used as definitive evidence of cheating.

Furthermore, the transition has exposed a lack of confidence in current technological solutions. Gamage et al. (2022) used interviews to investigate staff and student perceptions of the future of examinations. Interestingly, while both groups appreciate the utility of proctoring tools, there remains a strong preference for a return to paper-based exams due to a lack of technological self-efficacy. The study also raises a critical pedagogical concern: many instructors designing "innovative" online assessments have never experienced remote testing from a student perspective. This disconnect raises questions regarding how effectively these instructors can calibrate benchmarks or assess whether their learning outcomes are truly achievable in a digital environment.

### **Technology for Cheating Prevention and Detection**

There is a growing body of research regarding the design, development, implementation, and testing of innovative technologies to ensure the integrity of online exams. For example, the study of Erdem and Karabatak (2025) aims to develop deep learning and machine learning algorithms capable of identifying unethical behavior patterns within online exam data. By comparing synchronous data (collected in controlled computer labs) with asynchronous data, the researchers evaluated the performance of various models in detecting misconduct.

Similarly, Duham et al. (2021) propose a data-mining technique for detection; however, their approach focuses primarily on identifying similarities between student submissions and does not account for the specific nuances of AI-assisted cheating.

The broader landscape of these technologies is reviewed by Ustun et al. (2026), who provide a systematic literature survey on AI-supported proctoring. They categorize security measures into four distinct domains: authentication mechanisms, real-time monitoring, device access restriction, and general exam safety (p. 11). Notably, they include high-quality exam design and adaptive testing methods as essential components of the "exam safety" category.

The ethical implications of these tools are scrutinized by Coghlan et al. (2021), who examine the ethics of online supervision technologies. Their analysis, however, focuses exclusively on proctoring software and automated surveillance, omitting a discussion on the role of exam content or question design in maintaining integrity.

The reliability of such automated systems is questioned by Tweissi et al. (2022), who manually audited AI-based proctoring software for "false flags." Their findings reveal a significant discrepancy between human judgment and software-driven decisions; the system generated 74 incorrect flags across 244 exam attempts. This indicates that current proctoring capabilities require substantial refinement to reach human-level accuracy. Despite these limitations, development continues, as seen in the work of Jia and He (2022), who designed an intelligent proctoring system that utilizes AI-driven video and audio analysis to monitor for inappropriate behavior.

### **Method**

The present study utilized a controlled experimental design to investigate the intersection of student performance and artificial intelligence. The participant pool consisted of 15 freshmen recruited from the College of Business, ensuring a homogeneous sample in terms of academic standing and disciplinary focus.

The experimental assessment was structured to evaluate the divergent capabilities of human cognition and artificial intelligence across varying levels of contextual specificity. The assessment comprised ten unique questions, categorized based on the expected knowledge gaps of the participants and the AI tool.

The first six questions were sourced from a test bank accompanying a senior-level Operations Management textbook. Given that the participants were freshmen, it was hypothesized that the students would lack the specialized knowledge to answer these independently, whereas the AI would possess the requisite information. Within this subset, questions 1 through 5 were information-based, focusing on specific chapter content, while the sixth question required analysis of a specialized case study presented in the text.

In contrast, questions 7, 8, and 10 were original items developed by the researchers. These questions focused on specific topics discussed during lectures and incorporated elements of the students' unique, in-class personal experiences. The researchers anticipated that the students would be able to answer these correctly based on their lived participation, while the AI—lacking access to the specific classroom context—would not. Question 9 served as a baseline generic item regarding course topics, which both the students and the AI were expected to answer successfully.

To isolate the impact of generative tools, the study employed a within-subjects design. Participants were required to engage with each question in two distinct phases:

- **Iteration I (Independent):** Students completed the assessment relying exclusively on their prior knowledge without external aids.
- **Iteration II (Assisted):** Students were permitted to re-attempt the same questions utilizing AI assistance.

This dual-iteration approach allowed for a comparative analysis of how AI integration influences performance across different question types, ranging from standardized textbook material to highly contextualized classroom experiences.

To maintain the integrity and reliability of the experimental environment, the entire process was strictly proctored and supervised by the research team. This oversight ensured that the transition between independent and assisted phases was clearly delineated and that the use of AI was consistent with the study's protocols.

Complementing the quantitative assessment data, students participated in a qualitative survey administered both prior to and following the quiz. This instrument featured three open-ended questions designed to gauge the students' experiences, perceptions, and evolving attitudes toward AI integration in academic tasks. This dual-

method approach allowed for a comprehensive analysis of both performance metrics and student-reported insights.

## Findings

To evaluate the statistical significance of the performance gap between student and AI-generated responses, a two-sample t-test assuming unequal variances was conducted. The results for the initial five questions are summarized in Table 1. Given that students were not expected to possess prior knowledge of these specific items, any correct responses in the independent phase were attributed to random chance. The analysis reveals that for Questions 2, 3, and 4, generative AI significantly outperformed the students ( $p < 0.05$ ), aligning with the initial hypothesis. Conversely, for Question 1, student performance via random selection was statistically comparable to the AI-generated responses. Notably, in Question 5, the AI's performance was inferior to that of the students, failing to demonstrate the expected technical superiority for this item.

Table 1. t-test Results for Q1 to Q5

|                  | Q1          |      | Q2          |      | Q3          |      | Q4          |      | Q5          |       |
|------------------|-------------|------|-------------|------|-------------|------|-------------|------|-------------|-------|
|                  | Studen      | GenA | Studen      | GenA | Studen      | GenA | Studen      | GenA | Studen      | GenAI |
|                  | t           | I    | t           | I    | t           | I    | t           | I    | t           |       |
| Mean             | 0.80        | 0.87 | 0.13        | 0.93 | 0.40        | 1.00 | 0.47        | 0.93 | 0.27        | 0.00  |
| Variance         | 0.17        | 0.12 | 0.12        | 0.07 | 0.26        | 0.00 | 0.27        | 0.07 | 0.21        | 0.00  |
| Observations     | 15          | 15   | 15          | 15   | 15          | 15   | 15          | 15   | 15          | 15    |
| Hypothesized     | 0           |      | 0           |      | 0           |      | 0           |      | 0           |       |
| Mean             |             |      |             |      |             |      |             |      |             |       |
| Difference       |             |      |             |      |             |      |             |      |             |       |
| df               | 27          |      | 26          |      | 14          |      | 21          |      | 14          |       |
| t Stat           | -0.48       |      | -7.10       |      | -4.58       |      | -3.13       |      | 2.26        |       |
| P(T<=t) two-tail | <b>0.64</b> |      | <b>0.00</b> |      | <b>0.00</b> |      | <b>0.01</b> |      | <b>0.04</b> |       |
| t Critical       | 2.05        |      | 2.06        |      | 2.14        |      | 2.08        |      | 2.14        |       |
| two-tail         |             |      |             |      |             |      |             |      |             |       |

Table 2 presents the t-test results for Question 6. Although this item was sourced from the same textbook test bank as the preceding questions, it required an analysis of a specific case study rather than the general theoretical topics covered in the chapters. The statistical analysis indicates that the generative AI failed to identify the correct response. Consequently, there was no significant difference between the AI-generated outputs and the responses provided by students via random selection. This outcome suggests a limitation in the AI's ability to process or access specific, isolated case-study narratives compared to broader thematic content.

Table 3 summarizes the t-test results for Questions 7, 8, and 10. These items were specifically designed to assess knowledge accessible to the students through their unique classroom experiences but presumed to be outside the AI's training data. The analysis revealed no significant difference between student and AI-generated responses. In interpreting these results, correct AI responses are attributed to statistical chance, while incorrect student

responses are categorized as standard performance errors inherent to any examination environment.

Table 2. t-test Results for Q6

|                              | Q6          |       |
|------------------------------|-------------|-------|
|                              | Student     | GenAI |
| Mean                         | 0.07        | 0.07  |
| Variance                     | 0.07        | 0.07  |
| Observations                 | 15          | 15    |
| Hypothesized Mean Difference | 0           |       |
| df                           | 28          |       |
| t Stat                       | 0           |       |
| P(T<=t) two-tail             | <b>1.00</b> |       |
| t Critical two-tail          | 2.05        |       |

Table 3. t-test Results for Q7, Q8 and Q10

|                              | Q1          |       | Q2          |       | Q3          |       |
|------------------------------|-------------|-------|-------------|-------|-------------|-------|
|                              | Student     | GenAI | Student     | GenAI | Student     | GenAI |
| Mean                         | 0.87        | 1.00  | 0.60        | 0.53  | 0.60        | 0.73  |
| Variance                     | 0.12        | 0.00  | 0.26        | 0.27  | 0.26        | 0.21  |
| Observations                 | 15          | 15    | 15          | 15    | 15          | 15    |
| Hypothesized Mean Difference | 0           |       | 0           |       | 0           |       |
| df                           | 14          |       | 28          |       | 28          |       |
| t Stat                       | 1.76        |       | 1.70        |       | 1.70        |       |
| P(T<=t) two-tail             | <b>0.16</b> |       | <b>0.72</b> |       | <b>0.46</b> |       |
| t Critical two-tail          | 2.14        |       | 2.05        |       | 2.05        |       |

Finally, Table 4 presents the t-test results for Question 9, which served as a baseline item that both the students and the generative AI were expected to answer correctly.

Table 4. t-test Results for Q9

|                              | Q6          |       |
|------------------------------|-------------|-------|
|                              | Student     | GenAI |
| Mean                         | 0.07        | 1     |
| Variance                     | 0.24        | 0     |
| Observations                 | 15          | 15    |
| Hypothesized Mean Difference | 0           |       |
| df                           | 14          |       |
| t Stat                       | -2.650      |       |
| P(T<=t) two-tail             | <b>0.02</b> |       |
| t Critical two-tail          | 2.15        |       |

As hypothesized, the AI achieved universal accuracy, providing the correct response in every instance. In contrast, the student performance reflected a degree of variance; the observed errors are attributed to standard performance errors typical of a standard examination environment.

### Qualitative Analysis

Prior to and following the experimental phase, participants completed a qualitative survey consisting of three open-ended questions designed to capture their perceptions and feedback. The resulting qualitative data were analyzed in conjunction with the researchers' direct observations recorded during the assessment. Through a thematic analysis of these combined data sources, three primary recurring themes were identified, providing a comprehensive view of the students' evolving engagement with the AI-assisted environment.

The first emergent theme centered on a pronounced over-reliance on the generative model; several participants demonstrated a lack of critical engagement with the AI-generated outputs. These students frequently prioritized the speed of completion over comprehension, directly copying responses before the model could provide the underlying rationale or explanation. One participant characterized this tendency, stating:

*"[I used ChatGPT] because... it gives a quick response."*

Whereas another student reported:

*"I like ChatGPT because I have a shortcut... [that] helps me get information into an easier to read version."*

Another significant consideration involves the varying levels of participant diligence, particularly given that the assessment was not graded. While a subset of students demonstrated high levels of engagement, dedicating ample time to the task and meticulously reviewing their submissions, others exhibited less rigor. Notably, certain participants revisited their initial, independent responses after viewing the AI-generated outputs. Furthermore, several students returned to previous items to evaluate their work and occasionally modified their answers based on subsequent insights.

The final theme emerged from the students' critical evaluation of the generative model's reliability. Participants demonstrated varying degrees of skepticism; while some meticulously audited the AI's outputs and identified significant inaccuracies, others were impressed by the model's capabilities. Specifically, some students noted that the AI successfully addressed items related to localized classroom discussions. Despite these instances of success, several participants offered pointed critiques of the AI's performance, particularly regarding its inconsistency. Some students remarked:

*"[It's] interesting how AI also fails certain questions... we cannot rely on it for the whole future."*

*“[I was] surprised by how many questions AI got wrong.”*

whereas some students were surprised by its unexpected performance:

*“Most of them AI got right, even questions that we covered only in class.”*

## Conclusion

The results from this study show how important it is to change how we write tests in the age of AI. The main lesson is that the way questions are built matters. When teachers move away from generic questions and instead ask about specific things that happened in class, AI becomes much less helpful. This approach not only makes it harder to cheat but also encourages students to pay more attention and participate during lectures.

However, the study also reveals a major challenge: AI models are becoming so smart and complex that it is getting harder to stay one step ahead of them. Designing "AI-proof" questions takes a lot of effort and a deep understanding of what the AI does and does not know. This suggests that simply changing test questions might not be enough on its own to solve the problem of academic dishonesty.

Finally, while these results are a good start, they are based on a small group of students. More research is needed with larger groups of students from different majors to see if these strategies work the same way in every subject. This will help universities find a fair and lasting balance between using new technology and making sure students are actually learning.

## Statements and Declarations

**Acknowledgments and Notes:** This research is supported by Alfred University’s Bernstein Funds for Student and Faculty Development, and Shaner College of Business Support Fund, therefore, the authors would like to thank the Bernstein Family and Lance T. Shaner for their generous contributions.

**AI Use:** Not applicable.

**Supplementary Materials:** Not applicable.

**Author Contributions:** Not applicable.

**Funding:** Not applicable.

**Data Availability:** Not applicable.

**Ethics Approval:** Not applicable.

**Informed Consent:** Not applicable.

**Conflicts of Interest:** Not applicable.

## References

- Bennett, S. (2026). *Academic integrity in Massive Open Online Courses (MOOCs): Scale challenges and solutions for 2026*. Paper Checker
- Bittle, K., & El-Gayar, O. (2025). Generative AI and academic integrity in higher education: A systematic review and research agenda. *Information*, 16(4), 296. <https://doi.org/10.3390/info16040296>
- Cantiello, J., & Geschke, R. H. (2024). Preventing academic dishonesty in online courses: Best practices to discourage cheating. *Journal of Health Administration Education*, 40(2), 205–230.
- Coghlan, S., Miller, T., & Paterson, J. (2021). Good proctor or “Big Brother”? *Ethics of online exam supervision technologies*. *Philosophy & Technology*, 34, 1581–1606. <https://doi.org/10.1007/s13347-021-00476-1>
- Duhaim, A. M., Al-Mamory, S., & Mahdi, M. S. (2022). Cheating detection in online exams during COVID-19 pandemic using data mining techniques. *Webology*, 19(1), 341–366. <https://doi.org/10.14704/WEB/V19I1/WEB19026>
- Erdem, B., & Karabatak, M. (2025). Cheating detection in online exams using deep learning and machine learning. *Applied Sciences*, 15(1), 400. <https://doi.org/10.3390/app15010400>
- Evangelista, E. D. L. (2025). Ensuring academic integrity in the age of ChatGPT: Rethinking exam design, assessment strategies, and ethical AI policies in higher education. *Contemporary Educational Technology*, 17(1), ep559. <https://doi.org/10.30935/cedtech/15775>
- Gamage, K. A. A., Pradeep, R. G. G. R., & de Silva, E. K. (2022). Rethinking assessment: The future of examinations in higher education. *Sustainability*, 14(6), 3552. <https://doi.org/10.3390/su14063552>
- Gundu, T. (2025). Strategies for e-assessments in the era of generative artificial intelligence. *Electronic Journal of e-Learning*, 22(7), 40–50. <https://doi.org/10.34190/ejel.22.7.3477>
- Hanh, N. V., & Duyen, N. T. (2025). AI-assisted academic cheating: A conceptual model based on postgraduate student voices. *Frontiers in Computer Science*, 7, 1682190. <https://doi.org/10.3389/fcomp.2025.1682190>
- Hartley, P., Torres, L., & Routon, P. W. (2025). Academic integrity and exam performance in the online environment. *Journal of Business and Economic Studies*, 29(1), 64–78.
- Holden, O. L., Norris, M. E., & Kuhlmeier, V. A. (2021). Academic integrity in online assessment: A research review. *Frontiers in Education*, 6, 639814. <https://doi.org/10.3389/educ.2021.639814>
- Jia, J., & He, Y. (2022). The design, implementation and pilot application of an intelligent online proctoring system for online exams. *Interactive Technology and Smart Education*, 19(1), 112–120. <https://doi.org/10.1108/ITSE-12-2020-0246>
- Kara-Mohamed, M. (2025). The use of VLE assessments in the presence of AI technology. *Journal of Educational Technology Systems*, 54(1), 67–87. <https://doi.org/10.1177/00472395251351638>

- Martin, F., Kim, S., Bolliger, D. U., et al. (2025). Assessment types, strategies, and feedback in online higher education courses in the age of artificial intelligence: Perspectives of instructional designers. *TechTrends*, 69, 1330–1346. <https://doi.org/10.1007/s11528-025-01115-8>
- Nikolic, S., Daniel, S., Haque, R., Belkina, M., Hassan, G. M., Grundy, S., Lyden, S., Neal, P., & Sandison, C. (2023). ChatGPT versus engineering education assessment: A multidisciplinary and multi-institutional benchmarking and analysis of this generative artificial intelligence tool to investigate assessment integrity. *European Journal of Engineering Education*, 48(4), 559–614. <https://doi.org/10.1080/03043797.2023.2213169>
- Noorbehbahani, F., Mohammadi, A., & Aminazadeh, M. (2022). A systematic review of research on cheating in online exams from 2010 to 2021. *Education and Information Technologies*, 27, 8413–8460. <https://doi.org/10.1007/s10639-022-10927-7>
- Tweissi, A., Al Etaiwi, W., & Al Eisawi, D. (2022). The accuracy of AI-based automatic proctoring in online exams. *The Electronic Journal of e-Learning*, 20(4), 419–435.
- Uça Güneş, E. P., Gedik, N., Işıkoğlu, M. A., Yiğit, B., Güneş, İ., & Beylik, A. (2024). A case on remote supervised exam practices during the pandemic. *Open Praxis*, 16(4), 712–730. <https://doi.org/10.55982/openpraxis.16.4.691>
- Ustun, A. G., Kayali, B., Yavuz, M., & Bozkurt, A. (2026). Surveillance or support? A systematic literature review of artificial intelligence supported proctored online exams. *Turkish Online Journal of Distance Education*, 27(2), 1–24. <https://doi.org/10.17718/tojde.1674490>



## GenAI-Assisted Cybersecurity Mesh for Secure Digital Transactions

Utham Kumar Anugula Sethupathy <sup>1</sup>, Diwakar Sreeramulu <sup>2</sup>

<sup>1</sup> Independent Researcher & Industry Practitioner, AI and Cybersecurity, USA, 0009-0004-0661-043X

<sup>2</sup> Independent Researcher & Industry Practitioner, AI and Cybersecurity, USA,

\* Corresponding author: Utham Kumar Anugula Sethupathy (ilsetoffice@gmail.com)

### Article Info

#### Article History

Received:

1 February 2026

Revised:

25 April 2026

Accepted:

11 July 2026

Published:

30 August 2026

#### Keywords

Generative artificial intelligence  
Cybersecurity mesh architecture  
Digital transactions  
Distributed security systems  
Real-time risk detection

### Abstract

Digital transaction ecosystems increasingly operate across distributed platforms, heterogeneous services, and high velocity data flows. Traditional centralized and perimeter centric security models struggle to provide timely threat detection, adaptive response, and scalable enforcement under these conditions. This paper proposes a Generative Artificial Intelligence (GenAI)-assisted cybersecurity mesh architecture designed to secure digital transactions through decentralized, autonomous, and cooperative security agents. The architecture distributes intelligence across transaction entry points, application services, and data layers, enabling localized risk assessment while maintaining global situational awareness through shared threat intelligence. A detailed engineering framework is presented, including system components, agent orchestration workflows, and integration patterns with existing transaction processing pipelines. Experimental evaluation using simulated high throughput transaction workloads demonstrates improvements in detection latency, false positive reduction, and horizontal scalability when compared with monolithic security approaches. The findings indicate that GenAI assisted cybersecurity mesh architectures provide a practical and resilient foundation for securing modern digital transaction platforms. Modern digital payment platforms are increasingly exposed to fraud, automated abuse, and coordinated cyberattacks as transaction velocity and platform interconnectivity continue to grow (ENISA, 2023; Kshetri, 2021).

**Citation:** Sethupathy, U. K. A., & Sreeramulu, D. (2026). GenAI-Assisted Cybersecurity Mesh for Secure Digital Transactions. In M. Shelley, S. Turgut, & E. Cakir (Eds.), *Proceedings of International Conference on Life Sciences, Engineering and Technology* (pp. 12–24), New York, USA. ISTES.



Volume: 7, © 2026, Published by ISTES, ISBN: 978-1-952092-79-4, ISSN: 2833-0447



## Introduction

Digital payment and transaction platforms have evolved into complex, distributed ecosystems spanning cloud services, edge devices, application programming interfaces (APIs), and third-party integrations. These environments process millions of transactions per second while simultaneously facing sophisticated threats such as fraud, account takeover, automated abuse, and coordinated cyberattacks. Conventional security architectures, which rely heavily on centralized controls and static rule sets, often introduce latency and fail to adapt to rapidly changing threat patterns.

Recent advances in artificial intelligence have enabled more adaptive security mechanisms; however, many implementations remain constrained by centralized decision engines and limited contextual awareness. This paper argues that securing modern transaction systems requires both distributed enforcement and intelligent adaptation. To address this need, we propose a GenAI-assisted cybersecurity mesh architecture that decentralizes security intelligence while preserving coordinated response capabilities across the transaction ecosystem.

## Background and Related Work

### Transaction Security Architectures

Digital transaction security has historically relied on centralized fraud engines, rule-based filtering, and perimeter-centric defenses. These approaches were effective in early card-present and batch-processing environments but have shown structural limitations as transaction ecosystems evolved toward real-time, API-driven, and cloud-native architectures. Centralized architectures also introduce structural bottlenecks and single points of failure that limit scalability and resilience in distributed transaction systems (Kahn et al., 2020). Prior studies on transaction fraud detection highlight the limitations of centralized decision engines in adapting to evolving behavioral patterns and high-volume transaction streams (Li et al., 2022; Zhang et al., 2023).

### Cybersecurity Mesh Architecture

Cybersecurity Mesh Architecture (CSMA) has been proposed as a distributed security paradigm that shifts enforcement closer to protected assets while maintaining centralized policy intent (Gartner, 2022). CSMA emphasizes distributed enforcement, identity-centric access, and policy consistency across heterogeneous environments. Prior studies highlight improved resilience and scalability when security controls are placed closer to protected assets. Practical CSMA adoption requires careful coordination of identity, policy enforcement, and interoperability across heterogeneous environments (Hoover et al., 2022). Such distributed security models must also align with established information security management standards to ensure consistency and auditability (ISO/IEC, 2022).

### Generative AI in Security Systems

Advances in generative artificial intelligence have enabled models capable of learning complex representations

and reasoning over high-dimensional behavioral data (Goodfellow et al., 2016). In cybersecurity contexts, GenAI has been explored for log analysis, anomaly explanation, and adversarial simulation. Unlike discriminative classifiers, generative models can infer intent and generate adaptive responses, making them suitable for dynamic security environments. The integration of GenAI into distributed security architectures remains underexplored, particularly in high-throughput transaction systems. However, operational machine learning systems introduce hidden technical debt related to data dependencies, model drift, and deployment complexity (Sculley, D., Holt, G., Golovin, D., et al., 2015). Prior research has also demonstrated that learned models may exhibit unexpected behavior under adversarial or distribution-shifted inputs (Szegedy, C., Zaremba, W., Sutskever, I., et al., 2014).

## Proposed GenAI-Assisted Cybersecurity Mesh Architecture

### *System Design Objectives*

The proposed architecture is designed to satisfy five primary objectives: low-latency threat detection, horizontal scalability, resilience to partial failures, adaptive risk reasoning, and policy consistency across distributed transaction environments. These objectives reflect the operational constraints of modern digital payment systems, where security decisions must be executed within strict time budgets without degrading user experience or transaction throughput.

### *Architectural Overview*

The GenAI-assisted cybersecurity mesh distributes intelligent security agents across transaction ingress points, application services, and data persistence layers. Each agent operates autonomously while participating in a cooperative mesh that maintains shared situational awareness. Unlike centralized security engines, the mesh architecture eliminates reliance on a single control plane, thereby reducing decision latency and improving fault tolerance under peak load conditions. The proposed design extends cybersecurity mesh principles by embedding GenAI-based reasoning within distributed enforcement points, rather than relying on centralized analytic engines (Gartner, 2022). Table 1 summarizes the core components of the GenAI-assisted cybersecurity mesh and their respective security responsibilities.

Table 1. Cybersecurity Mesh Components

| Component                    | Primary Function               | Security Contribution       |
|------------------------------|--------------------------------|-----------------------------|
| Transaction Monitoring Agent | Inline transaction observation | Real-time anomaly detection |
| GenAI Reasoning Module       | Probabilistic risk inference   | Adaptive threat recognition |
| Policy Orchestration Layer   | Global intent distribution     | Consistent enforcement      |
| Threat Intelligence Mesh     | Agent coordination             | Rapid threat propagation    |

The architecture is logically divided into a data plane and a control plane. The data plane is responsible for real-time transaction observation, feature extraction, and localized enforcement actions. The control plane manages policy intent, trust calibration, and cross-agent coordination through secure message exchanges. This separation enables localized decision-making while preserving centralized governance. Figure 1 illustrates the proposed

GenAI-assisted cybersecurity mesh architecture, highlighting distributed agent placement, control and data plane separation, and threat intelligence propagation across transaction services.

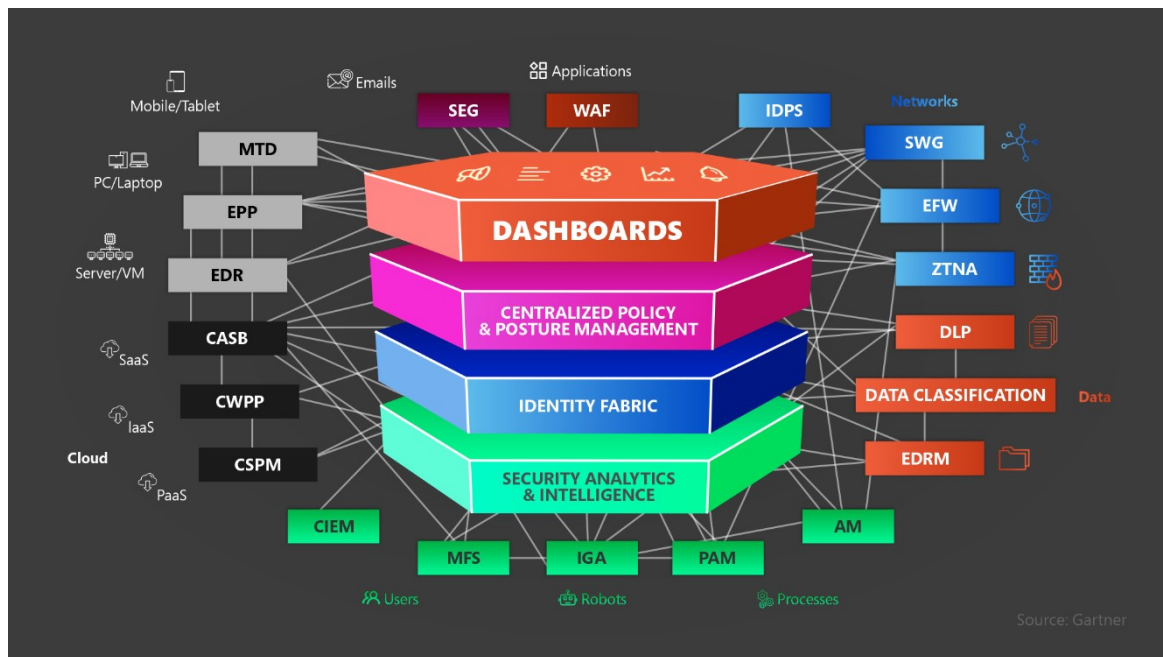


Figure 1. GenAI-Assisted Cybersecurity Mesh Architecture

### Core Architectural Components

Transaction Monitoring Agents are deployed at points where transactions are initiated, transformed, or authorized. These agents collect behavioral signals, contextual metadata, device attributes, and historical interaction patterns. Monitoring occurs inline to ensure minimal latency overhead.

GenAI Reasoning Modules, embedded within each agent, perform probabilistic inference over transaction context. Rather than producing binary outcomes, the modules generate risk distributions and threat hypotheses that support adaptive enforcement strategies. This generative reasoning enables detection of previously unseen or evolving attack patterns.

Policy Orchestration Layer maintains global security intent, including risk thresholds, compliance constraints, and escalation rules. Policies are expressed declaratively and disseminated across the mesh, allowing agents to interpret and enforce them locally based on contextual risk.

Threat Intelligence Mesh provides secure, low-latency communication among agents. Indicators of compromise, inferred threat patterns, confidence scores, and enforcement outcomes are propagated across the mesh, enabling cooperative learning and rapid convergence on emerging threats.

### Trust Boundaries and Security Controls

Trust boundaries are explicitly defined between external clients, transaction services, internal processing components, and data stores. Each boundary is protected by one or more monitoring agents capable of independent enforcement. Mutual authentication, encrypted communication channels, and integrity validation are enforced for all inter-agent messaging to mitigate spoofing, poisoning, or replay attacks. Threat boundaries and attack surfaces are identified using established threat modeling principles to guide agent placement and enforcement decisions (Shostack, 2014).

### Operational Workflow

When a transaction event occurs, the local agent performs feature extraction and risk inference using its GenAI reasoning module. Transactions below the configured risk threshold proceed without interruption. Elevated-risk events trigger step-up controls such as additional authentication, rate limiting, or temporary transaction holds. High-confidence threat signals are broadcast across the mesh, allowing peer agents to preemptively adjust enforcement policies.

This cooperative workflow enables rapid, context-aware response while avoiding cascading failures or unnecessary user friction.

## Methodology

### Threat Model and Assumptions

The evaluation assumes an adversarial environment in which attackers may attempt fraud, automated abuse, account takeover, and coordinated transaction manipulation. Adversaries are modeled as adaptive, capable of varying transaction velocity, payload characteristics, and behavioral signatures to evade static detection rules. Insider threats and infrastructure compromise are considered out of scope; however, the architecture assumes secure communication channels and authenticated agents within the mesh. The adversarial assumptions and attack scenarios are derived using structured threat modeling methodologies commonly applied in secure system design (Shostack, 2014). Table 2 defines the evaluation metrics used to assess performance, accuracy, scalability, and resilience of the proposed architecture.

Table 2. Evaluation Metrics Definition

| Metric              | Description                       | Measurement Objective  |
|---------------------|-----------------------------------|------------------------|
| Detection Latency   | Time to enforcement decision      | Real-time viability    |
| False-Positive Rate | Legitimate transactions flagged   | User experience        |
| Throughput          | Transactions processed per second | Scalability            |
| Resilience          | Performance under partial failure | Operational robustness |

The system operates under the assumption that transaction processing services expose sufficient telemetry, including request metadata, behavioral signals, and historical context, to support real-time risk inference. Latency constraints reflect industry-standard payment authorization time budgets. Figure 2 depicts the agent inference

lifecycle, including feature extraction, GenAI-based risk inference, enforcement selection, and cooperative intelligence sharing.

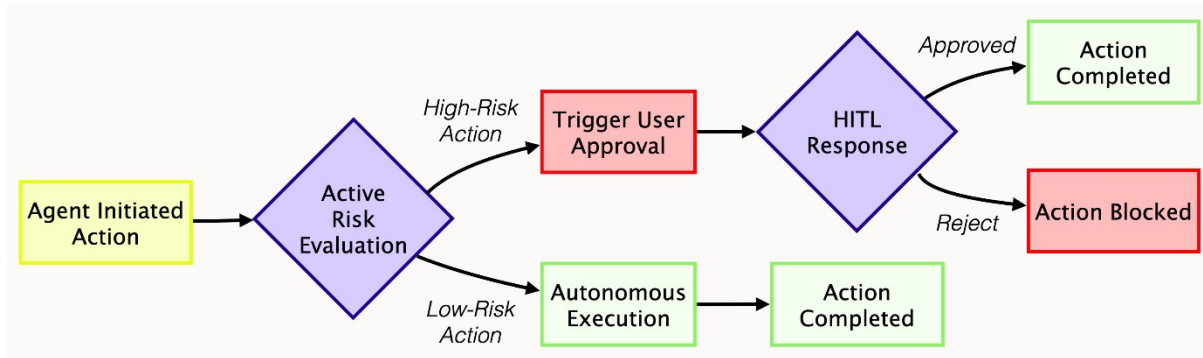


Figure 2. Agent Inference and Coordination Workflow

## Experimental Environment

A simulated transaction processing environment was constructed to evaluate the proposed architecture under controlled yet realistic conditions. The environment models high-velocity digital payment flows using synthetic transaction streams that vary in volume, concurrency, and threat composition. Transaction generators emulate legitimate user behavior alongside injected fraudulent and anomalous patterns, including burst attacks and low-and-slow evasion strategies.

The environment supports horizontal scaling of transaction monitoring agents and configurable deployment topologies to assess performance under increasing load. A centralized security baseline, representing a traditional monolithic fraud detection engine, is implemented for comparative evaluation.

## Agent Inference Lifecycle

Each transaction monitoring agent follows a structured inference lifecycle. Incoming transactions undergo feature extraction, including temporal patterns, behavioral deviations, and contextual attributes. The extracted features are processed by the embedded GenAI reasoning module, which produces a probabilistic risk score and one or more threat hypotheses. The generative reasoning process is grounded in deep learning architectures capable of modeling complex, non-linear behavioral relationships (Goodfellow et al., 2016).

Based on the inferred risk distribution, agents select enforcement actions aligned with global policy intent. Actions range from passive monitoring to step-up authentication or transaction suspension. High-confidence threat indicators are shared across the mesh to inform peer agents and support cooperative adaptation. To mitigate long-term degradation, the system accounts for known operational risks in machine learning deployments, including model drift and feedback loops (Sculley, D., Holt, G., Golovin, D., et al., 2015).

## Evaluation Metrics

System performance is evaluated using multiple quantitative metrics. Detection latency measures the elapsed time between transaction arrival and enforcement decision. False-positive rate captures the proportion of legitimate transactions incorrectly flagged as malicious. Throughput impact evaluates the effect of security processing on transaction completion rates. Scalability is assessed by measuring performance stability as agent instances and transaction volume increase.

## Baseline Comparison Strategy

The centralized baseline employs a single decision engine that aggregates transaction data and applies rule-based and statistical detection logic. Comparative analysis focuses on relative improvements in latency, accuracy, and scalability achieved by the GenAI-assisted cybersecurity mesh. Statistical significance is assessed through repeated simulation runs under identical workload conditions.

## Results

### Detection Latency Performance

The GenAI-assisted cybersecurity mesh demonstrates consistently lower detection latency compared to the centralized baseline across all evaluated workload conditions. Localized inference at transaction ingress points eliminates queuing delays associated with centralized decision engines. Under nominal transaction loads, average decision latency remains within strict authorization time budgets. As transaction volume increases, the latency curve for the centralized baseline exhibits superlinear growth, while the mesh architecture maintains near-constant response times. The observed performance trends are consistent with prior evaluations of distributed and adaptive fraud detection systems in financial environments (Li et al., 2022; Zhang et al., 2023). Table 3 provides a comparative summary of system performance, highlighting the advantages of the GenAI-assisted cybersecurity mesh over centralized security architectures.

Table 3. Comparative Performance Summary

| Dimension         | Centralized Model   | GenAI Mesh  |
|-------------------|---------------------|-------------|
| Detection Latency | Degrades under load | Stable      |
| False Positives   | Elevated            | Reduced     |
| Scalability       | Limited             | Near-linear |
| Failure Impact    | System-wide         | Localized   |

Under burst-traffic scenarios simulating coordinated fraud attempts, the centralized model experiences latency spikes due to contention at the decision engine. In contrast, the distributed agents absorb traffic surges independently, preventing congestion propagation. These results confirm that decentralizing inference is critical for maintaining real-time responsiveness in high-velocity transaction environments. As shown in Figure 3, the GenAI-assisted mesh maintains stable detection latency under increasing transaction load, whereas the centralized

baseline exhibits rapid degradation.

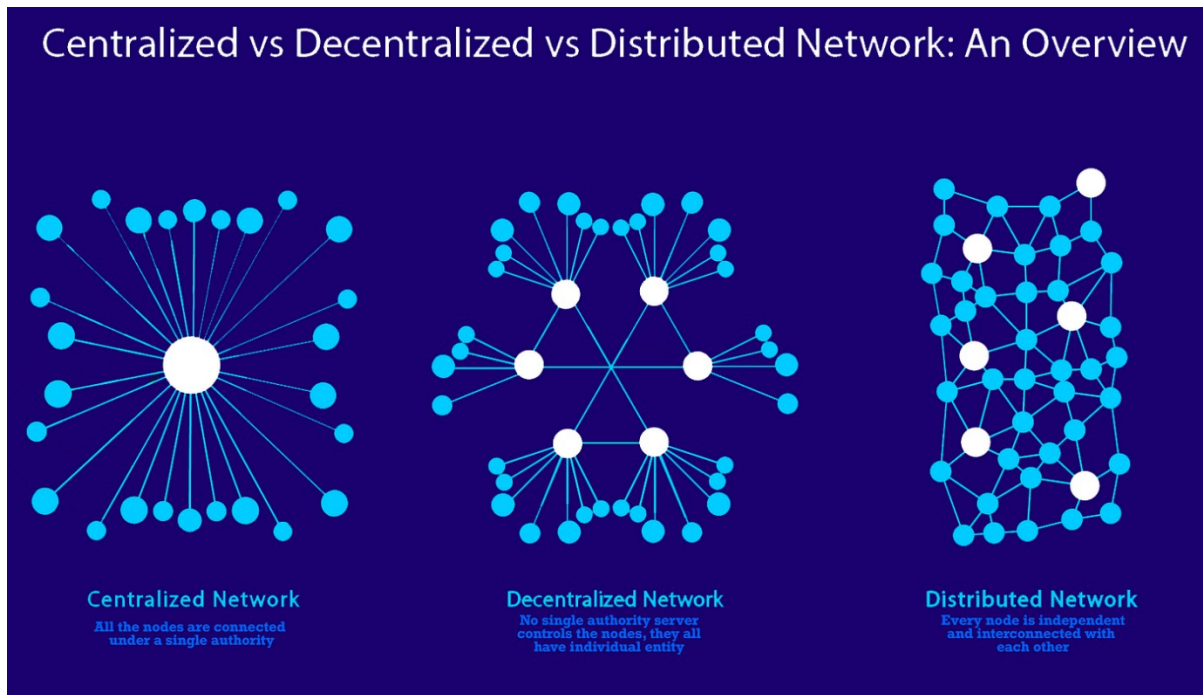


Figure 3. Detection Latency Comparison

### False-Positive Rate Analysis

False-positive rates are significantly reduced in the GenAI-assisted mesh relative to the centralized baseline. Local agents initially flag anomalous transactions based on contextual deviation; however, cooperative validation across the threat intelligence mesh enables rapid reconciliation of local observations with global behavioral patterns. Transactions that appear suspicious in isolation but align with legitimate cross-channel behavior are cleared without user-facing friction.

This cooperative reasoning mechanism reduces unnecessary transaction challenges and declines, improving customer experience while preserving security posture. The centralized baseline lacks this multi-perspective validation capability, leading to higher false-positive rates, particularly during periods of behavioral drift such as seasonal spending changes or promotional events. Figure 4 demonstrates a measurable reduction in false-positive rates enabled by cooperative risk validation across distributed security agents.

### Throughput and Scalability Evaluation

Scalability analysis demonstrates near-linear throughput gains as transaction monitoring agents are horizontally scaled. Adding agents proportionally increases processing capacity without degrading detection accuracy or latency. The absence of a centralized control bottleneck enables the system to maintain stable performance under increasing transaction volumes. These findings align with established principles of distributed systems, where decentralized processing improves throughput and fault tolerance under load (Kahn et al., 2020).

In contrast, the centralized baseline exhibits diminishing throughput returns beyond a moderate load threshold. As transaction volume increases, contention for shared resources limits effective scaling, resulting in delayed decisions and degraded user experience. These findings highlight the suitability of the proposed architecture for large-scale payment platforms with variable and unpredictable workloads.

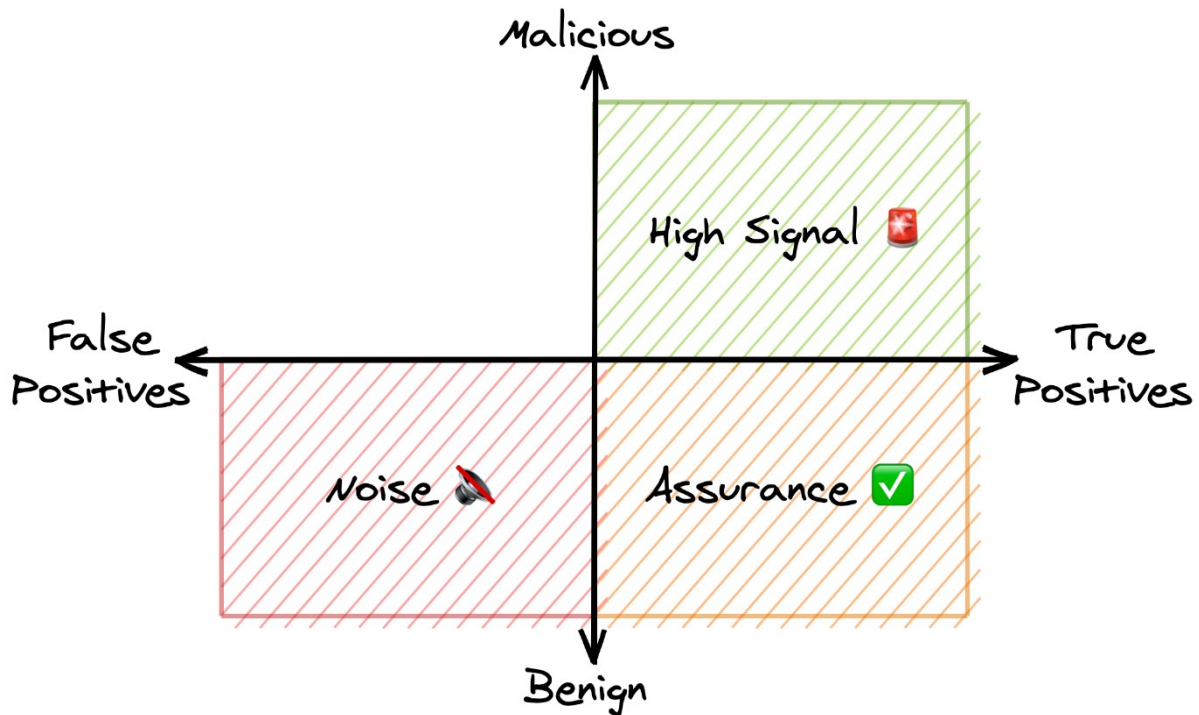


Figure 4. False-Positive Rate Reduction

### Resilience Under Partial Failure

The distributed nature of the cybersecurity mesh enhances resilience under partial failure conditions. Simulated agent failures do not result in systemic degradation, as neighboring agents dynamically compensate by adjusting enforcement sensitivity and intelligence sharing. Threat detection capability degrades gracefully rather than catastrophically.

Centralized architectures exhibit greater fragility under comparable failure scenarios. Loss of availability or performance degradation at the central decision engine impacts the entire transaction ecosystem. The mesh architecture's fault isolation properties therefore represent a significant operational advantage for mission-critical financial systems.

### Summary of Comparative Performance

Across all evaluation dimensions—latency, false-positive rate, throughput, scalability, and resilience—the GenAI-assisted cybersecurity mesh outperforms the centralized security baseline. The results indicate that

distributing GenAI-powered reasoning across cooperative agents enables faster, more accurate, and more robust security enforcement without compromising transaction performance.

## **Discussion**

### **Architectural Implications for Transaction Security**

The results demonstrate that distributing GenAI-powered reasoning across a cybersecurity mesh fundamentally alters how transaction security systems scale and respond to threats. Unlike centralized architectures that prioritize global visibility at the cost of latency, the mesh model enables localized intelligence while preserving coordinated enforcement. This shift aligns security decision-making with the distributed nature of modern transaction platforms, where risk signals are inherently fragmented across services, channels, and time.

The architecture shows that real-time security does not require centralized control, but rather coordinated autonomy. By embedding reasoning capabilities close to transaction execution points, the system reduces decision latency while maintaining contextual richness through shared intelligence. This finding challenges prevailing assumptions that effective fraud and abuse detection must rely on centralized aggregation.

### **Explainability and Trust Calibration**

A critical consideration in deploying GenAI-assisted security systems is explainability. Generative models inherently produce probabilistic and inferential outputs, which may be perceived as opaque by operators, auditors, and regulators. In the proposed architecture, explainability is addressed at the agent level through the generation of intermediate risk factors and threat hypotheses, rather than opaque binary decisions.

Trust calibration across agents is equally important. Overly aggressive propagation of threat signals can amplify false positives, while overly conservative sharing can delay response to coordinated attacks. The mesh architecture supports dynamic trust weighting, allowing agents to adjust reliance on peer intelligence based on confidence scores, historical accuracy, and contextual similarity. This adaptive trust model is essential for maintaining balance between responsiveness and stability.

### **Governance and Policy Management**

Decentralized intelligence introduces new governance challenges. While policy enforcement occurs locally, policy intent must remain globally consistent to ensure compliance and operational predictability. The separation of control plane and data plane plays a central role in addressing this challenge. Declarative policy definitions enable centralized oversight without reintroducing centralized decision bottlenecks. Emerging AI governance frameworks emphasize the importance of transparency, risk management, and accountability in AI-assisted decision systems (NIST, 2023).

Operational governance also extends to lifecycle management of GenAI models. Model updates, retraining

schedules, and drift detection must be coordinated across agents to prevent inconsistent behavior. The mesh architecture supports staged rollouts and selective activation, reducing systemic risk during updates.

### **Regulatory and Compliance Considerations**

Financial transaction systems operate under stringent regulatory frameworks that require auditability, consistency, and demonstrable risk controls. The proposed architecture aligns with these requirements by preserving traceable decision paths and enforcing policy constraints uniformly across agents. Distributed enforcement does not imply regulatory fragmentation; rather, it enables compliance controls to be applied closer to the point of risk manifestation. Alignment with established security and compliance standards remains essential for adoption in regulated financial environments (ISO/IEC, 2022; PCI SSC, 2022).

However, regulators may require additional assurance regarding the use of GenAI in security decision-making. Clear documentation of model behavior, decision boundaries, and fallback mechanisms is necessary to establish confidence. The architecture's ability to degrade gracefully and revert to conservative enforcement under uncertainty supports compliance expectations.

### **Operational Risk and Limitations**

While the cybersecurity mesh enhances resilience, it introduces operational complexity. Increased numbers of autonomous agents require robust observability, monitoring, and orchestration capabilities. Failure to manage these aspects may offset the benefits of decentralization.

Additionally, the evaluation is based on simulated transaction environments. Although the simulation captures realistic patterns and adversarial behavior, production deployments may surface edge cases related to data quality, integration constraints, and organizational readiness. These limitations underscore the need for phased adoption and continuous validation.

### **Implications for Future Transaction Platforms**

The findings suggest that GenAI-assisted cybersecurity mesh architectures are well suited for emerging transaction paradigms, including API-driven payments, embedded finance, and real-time digital wallets. As transaction ecosystems continue to decentralize, security architectures must evolve accordingly. The mesh model provides a scalable foundation that can adapt to future threat landscapes without sacrificing performance or governance.

### **Conclusion**

This paper presented a GenAI-assisted cybersecurity mesh architecture designed to address the scalability, latency, and adaptability challenges inherent in modern digital transaction systems. By distributing intelligent

security agents across transaction ingress points, application services, and data layers, the proposed architecture decentralizes decision-making while preserving coordinated enforcement through shared threat intelligence. This approach directly aligns security controls with the distributed nature of contemporary payment and transaction platforms.

The evaluation demonstrates that embedding GenAI-powered reasoning at the edge of transaction flows significantly improves detection latency, reduces false-positive rates, and enables near-linear scalability under increasing load. Compared to centralized security models, the mesh architecture exhibits superior resilience to burst traffic, partial system failures, and behavioral drift. These characteristics are critical for maintaining both security effectiveness and user experience in high-velocity financial environments.

Beyond performance gains, the study highlights important governance and compliance considerations. The separation of control and data planes enables consistent policy enforcement without reintroducing centralized bottlenecks, while agent-level explainability supports auditability and regulatory alignment. The findings suggest that decentralized intelligence, when coupled with coordinated policy oversight, can satisfy stringent financial compliance requirements while enabling adaptive security responses.

Future research will focus on validating the architecture in production-grade transaction environments, exploring advanced explainability techniques for generative reasoning, and integrating regulatory-aware policy frameworks. Additional work is needed to assess long-term model drift, cross-institution threat intelligence sharing, and the operational economics of large-scale deployment. Overall, GenAI-assisted cybersecurity mesh architectures represent a practical and forward-looking foundation for securing next-generation digital transaction ecosystems. As cybersecurity mesh architectures mature, alignment with evolving regulatory and security frameworks will be critical to sustaining trust and operational viability (Gartner, 2022; NIST, 2024).

## Statements and Declarations

**Acknowledgments and Notes:** Not applicable.

**AI Use:** Not applicable.

**Supplementary Materials:** Not applicable.

**Author Contributions:** Not applicable.

**Funding:** Not applicable.

**Data Availability:** Not applicable.

**Ethics Approval:** Not applicable.

**Informed Consent:** Not applicable.

**Conflicts of Interest:** Not applicable.

## References

- ENISA (2023). *Threat landscape for financial services*. European Union Agency for Cybersecurity. <https://www.enisa.europa.eu>
- Gartner (2022). *Cybersecurity mesh architecture (CSMA)*. Gartner Research.
- Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep learning*. MIT Press.
- Hoover, J., Shoard, P., & Gaehtgens, F. (2022). *How to start building a cybersecurity mesh architecture*. Gartner.
- ISO/IEC (2022). *ISO/IEC 27001: Information security management systems—Requirements*. International Organization for Standardization.
- Kahn, R., Rivest, R., & Shamir, A. (2020). Secure transaction processing in distributed systems. *IEEE Security & Privacy*, 18(4), 72–80. <https://doi.org/10.1109/MSEC.2020.3002054>
- Kshetri, N. (2021). *Cybersecurity and cyberwar: What everyone needs to know*. Oxford University Press.
- Li, Y., Xu, Y., & Chen, H. (2022). Distributed anomaly detection for financial transactions using machine learning. *IEEE Transactions on Services Computing*, 15(4), 2198–2210. <https://doi.org/10.1109/TSC.2020.3044821>
- NIST (2023). *AI risk management framework (AI RMF 1.0)*. National Institute of Standards and Technology. <https://www.nist.gov>
- NIST (2024). *Cybersecurity framework (CSF) 2.0*. National Institute of Standards and Technology. <https://www.nist.gov>
- Payment Card Industry Security Standards Council (2022). *Payment Card Industry Data Security Standard (PCI DSS) v4.0*. <https://www.pcisecuritystandards.org>
- Sculley, D., Holt, G., Golovin, D., et al. (2015). Hidden technical debt in machine learning systems. *Advances in Neural Information Processing Systems*, 28, 2503–2511.
- Shostack, A. (2014). *Threat modeling: Designing for security*. Wiley.
- Szegedy, C., Zaremba, W., Sutskever, I., et al. (2014). Intriguing properties of neural networks. *International Conference on Learning Representations (ICLR)*.
- Zhang, J., Luo, X., & Lee, J. (2023). Real-time fraud detection in payment systems using adaptive AI models. *ACM Transactions on Management Information Systems*, 14(2), 1–28. <https://doi.org/10.1145/3570912>



## A Qualitative Study of University Student Experiences Regarding the Impact of Artificial Intelligence on Coding Coursework

Yavuz Keceli <sup>1</sup>, Andrew W. Maguire <sup>2</sup>, Jack C. Meyer <sup>3\*</sup>

<sup>1</sup> Alfred University, USA, keceli@alfred.edu, 0000-0002-8981-8217

<sup>2</sup> Alfred University, USA, awm3@alfred.edu, 0009-0000-1260-2007

<sup>3</sup> Alfred University, USA, jcm17@alfred.edu, 0009-0009-7195-8340

\* Corresponding author: Jack C. Meyer (ilsetoffice@gmail.com)

### Article Info

#### Article History

Received:

24 February 2026

Revised:

25 April 2026

Accepted:

11 July 2026

Published:

30 August 2026

#### Keywords

Artificial intelligence

Higher education

Coding

Programming

### Abstract

The impact of generative artificial intelligence (AI) on coding and software development is an ongoing debate. As AI is changing how computer programs are created, its implications on college level coding and programming courses cannot be ignored by educators. Therefore, this study was conducted to assess the impact of artificial intelligence on coding coursework from the perspectives of university students. Throughout one semester, university students in a programming class were asked to complete several coding assignments individually, and several coding assignments using various generative AI tools. At the end of the semester, the students were asked to participate in a focus group study, comparing their manual coding output with code generated by AI, as well as sharing their perceptions of learning experiences. Based on the results, several recommendations were proposed for educators to improve their coding and programming courses to align with the evolving demands and the requirements of the job environment.

**Citation:** Keceli, Y., Maguire, A. W., & Meyer, J. C. (2026). A Qualitative Study of University Student Experiences Regarding the Impact of Artificial Intelligence on Coding Coursework. In M. Shelley, S. Turgut, & E. Cakir (Eds.), *Proceedings of International Conference on Life Sciences, Engineering and Technology* (pp. 25–30), New York, USA. ISTES.



Volume: 7, © 2026, Published by ISTES, ISBN: 978-1-952092-79-4, ISSN: 2833-0447



## Introduction

The rapid proliferation of generative artificial intelligence has fundamentally altered the landscape of computer science education, presenting both transformative opportunities and significant pedagogical challenges. While such tools offer the promise of increased efficiency and real-time support for new learners and novice programmers, their integration into the classroom has raised critical questions regarding the nature of student learning, critical thinking, and the authenticity of academic work. Current literature suggests a tension between the utility of AI for rapid problem-solving and the risk of fostering a crutch mentality that leads to over-reliance and an illusion of competence.

Therefore, this paper explores student perceptions of independent coding and AI-assisted programming. By examining qualitative data from undergraduate students engaged in a data analytics class using R programming, this study investigates the impact of student motivation and perceived learning outcomes by comparing their own assignments accomplished through both independent coding and artificial intelligence support.

## Literature Review

The use of generative AI is gaining significant attention in academic research, with a growing body of work examining its pedagogic impact on learning to code at the college level. Reihanian et al. (2025) provide a systematic literature review regarding the opportunities and challenges of using generative AI in coding education, specifically concerning accuracy, authenticity, and assessment. The authors argue that while AI tools offer exciting potential for enhancing efficiency and creativity in teaching foundational skills—such as coding, computational thinking, and problem-solving—the accuracy of AI-generated solutions remains a major concern. AI "hallucinations" can produce incorrect or irrelevant information, and students may lack the expertise to effectively fact-check these responses. Furthermore, while tools like ChatGPT and GitHub Copilot provide benefits such as real-time feedback and creative assistance, they introduce risks including error propagation and bias, which can adversely affect learning outcomes, particularly for novice learners.

Liffiton et al. (2023) express concern regarding students' over-reliance on large language models (LLMs) and propose an LLM-based tool designed to support learning without directly revealing solutions.

Similarly, Prather et al. (2024) discuss the benefits and harms of generative AI for novice programmers. They argue that although AI tools help students solve programming problems, the impact on learning remains largely unknown, especially when AI replaces, rather than supplements, critical thinking. Their research suggests that while generative AI accelerates problem-solving, inexperienced programmers may struggle to recognize their own lack of understanding due to metacognitive difficulties.

This finding aligns with Lee et al. (2025), who surveyed 319 knowledge workers to model how they enact critical thinking when using generative AI. Their results indicate that workers engage in critical thinking primarily to ensure work quality—such as verifying outputs against external sources—because AI can inhibit critical

engagement and lead to long-term over-reliance at the expense of independent problem-solving skills. Pădurean et al. (2025) similarly indicate that computing students often rely on generative AI without proper guidance, highlighting an urgent need to teach students how to interact with these tools effectively.

Hazzan and Erez (2025) argue that generative AI is disrupting computer science education (CSE) in multiple ways, suggesting that the community must embrace these technologies and reinvent its teaching methods and materials to remain relevant. Investigating this further, Jost et al. (2024) examined the impact of LLMs on programming education through an experimental study. They found that higher engagement with LLMs can detract from learning outcomes depending on the purpose of use; specifically, a significant inverse correlation exists between LLM use for code generation/debugging and the ability to solve tasks independently. Conversely, using LLMs to seek additional explanations showed no clear negative impact, suggesting such use may serve as a helpful supplement rather than a crutch.

Looking toward curriculum design, Munoz and Rook (2026) focus on simulation-based education using generative AI and "vibe coding." They describe vibe coding as "an informal mode of interacting with LLMs to produce code, one that tolerates imprecision in favor of creative momentum." They argue that educators can leverage this approach to create HTML and JavaScript-based simulations even without advanced coding skills, thereby reducing barriers to entry.

Finally, recent student-focused studies highlight the complexity of AI integration. Singer-Freeman et al. (2025) found that students generally consider using AI to create code acceptable unless instructed otherwise, reporting both positive perceptions (enhancing quality, compensating for weaknesses) and negative ones (interfering with learning, enabling grade-seeking without actual comprehension). Li et al. (2026) analyze student LLM usage patterns and advocate for a reform of traditional homework models. They suggest a shift toward diversified assessments that require students to interpret the internal logic of AI-generated code and identify its defects, moving the focus from writing code independently to critically evaluating and optimizing AI-generated content.

## Method

This study was conducted during a data analysis course involving R programming at a small liberal arts university in Upstate New York. Over the course of the semester, 24 students completed several coding assignments. In some of these tasks, students were instructed to input the assignment descriptors into generative AI tools of their choice to generate code. For these specific assignments, students were required to include the AI-generated code alongside their own reflections on its effectiveness and quality in their submissions.

Conversely, in other assignments, students were given similar tasks but were instructed to write the code independently, applying classroom learning without the assistance of AI. At the conclusion of the semester, students participated in focus groups to discuss their coding experiences, their level of effort, and their attitudes toward the impact of AI on their perceived learning. These focus groups were moderated by two research assistants trained in the methodology. Student responses to a set of predetermined questions were recorded and analyzed

qualitatively.

## Findings

Qualitative analysis of student responses revealed three recurring themes. The first theme concerns *prompt engineering*. Participants reported that AI tools were largely ineffective for coding coursework without extensive prompting to achieve desired outcomes. Students noted a constant need to redirect the AI; some even intentionally simplified AI outputs to appear more 'human.' One respondent noted, 'Artificial intelligence almost never does a good job the first time; I had to [repeatedly prompt] it into the right job,' while several others observed that AI frequently 'hallucinates' data to fit a specific context. This finding aligns with Liffiton et al. (2023), Pădurean et al. (2025), Jost et al. (2024), and Hazzan and Erez (2025), suggesting that prompt engineering is a critical skill. Consequently, educators may need to adapt curricula to ensure students acquire these competencies (Li et al., 2026; Murray et al., 2026).

The second theme involves *a diminished sense of connection to the work*. Using AI led students to feel less invested in their assignments and hindered their perceived learning. Participants remarked that 'AI made it harder to learn' and 'took away from being taught,' with others reporting a lack of motivation to exert effort on final projects. These sentiments reflect existing literature; for instance, Joshi and Vogel (2024) argue that AI-generated text decreases psychological ownership, though this effect is mitigated by longer, more involved prompts.

Similarly, Prather et al. (2024) suggest that while AI can accelerate progress for high-achieving students, it may leave struggling students with an 'illusion of learning.'

Finally, the third theme addresses *academic integrity and ethical perceptions*. Surprisingly, some students reported positive feelings regarding their use of AI, specifically stating they did not feel they were cheating. These participants viewed AI as a tool to 'fix models' rather than replace their own effort. This perspective supports the findings of Jost et al. (2024) and Prather et al. (2024), which advocate for AI as a supplementary rather than a substitutionary tool. In conclusion, students experience less guilt and maintain a stronger sense of ownership when AI usage is clearly defined and integrated into the curriculum as a supportive resource.

## Conclusion

This study highlights the complex impact of generative AI on student engagement in programming education. Our qualitative findings reveal that while students acknowledge the necessity of prompt engineering as a modern technical skill, they also experience a distinct psychological detachment from their work when AI facilitates the heavy lifting of code generation. This sense of diminished connection, coupled with concerns over learning inhibition, validates theoretical warnings regarding the potential for AI to undermine the grit and cognitive effort traditionally required to master programming.

However, the study also highlights a constructive pathway forward: when students utilize AI as a supplementary

tool for refinement rather than a substitutionary tool for creation, they maintain a sense of ethical integrity and intellectual ownership. Ultimately, the successful integration of AI into coding curricula depends on a shift in focus from mere code output to the critical evaluation of AI logic. By designing assessments that reward skilled and appropriate prompting and the optimization of generated content, educators can transform AI from a potential threat to academic rigor into a robust pedagogical partner that enhances, rather than replaces, student accomplishment.

## Statements and Declarations

**Acknowledgments and Notes:** This research is supported by Alfred University’s Bernstein Funds for Student and Faculty Development, and Shaner College of Business Support Fund, therefore, the authors would like to thank the Bernstein Family and Lance T. Shaner for their generous contributions.

**AI Use:** Not applicable.

**Supplementary Materials:** Not applicable.

**Author Contributions:** Not applicable.

**Funding:** Not applicable.

**Data Availability:** Not applicable.

**Ethics Approval:** Not applicable.

**Informed Consent:** Not applicable.

**Conflicts of Interest:** Not applicable.

## References

- Hazzan, O., & Erez, Y. (2025). Rethinking computer science education in the age of GenAI. *ACM Transactions on Computing Education*, 25(3), Article 26. <https://doi.org/10.1145/3732792>
- Joshi, N., & Vogel, D. (2024). Writing with AI lowers psychological ownership, but longer prompts can help. In *Proceedings of the 7th ACM Conference on Conversational User Interfaces*. Article 72, 1-17, <https://doi.org/10.1145/3719160.3736608>
- Jost, G., Taneski, V., & Karakatic, S. (2024). The impact of large language models on programming education and student learning outcomes. *Applied Sciences*, 14, 4115. <https://doi.org/10.3390/app14104115>

- Lee, H.-P., Sarkar, A., Tankelevitch, L., Drosos, I., Rintel, S., Banks, R., & Wilson, N. (2025). *The impact of generative AI on critical thinking: Self-reported reductions in cognitive effort and confidence effects from a survey of knowledge workers*. In Proceedings of the CHI Conference on Human Factors in Computing Systems (CHI '25). <https://doi.org/10.1145/3706598.3713778>
- Liffiton, M., Sheese, B., Savelka, J., & Denny, P. (2023). *Code help: Using large language models with guardrails for scalable support in programming classes*. In Proceedings of the 23rd Koli Calling International Conference on Computing Education Research (Koli Calling '23). <https://doi.org/10.1145/3631802.3631830>
- Li, H., Xiao, L., Lu, K., Li, D., Zhang, Z., & Xia, Q. (2026). Exploring the application of large language models (LLMs) in data structure instruction: An empirical analysis of student learning outcomes in computer science. *Information*, 17(4), 353. <https://doi.org/10.3390/info17040353>
- Munoz, A., & Rook, L. (2026). Leveraging generative AI through vibe coding: A case of simulation-based curriculum redesign in management education. *Education Sciences*, 16(4), 558. <https://doi.org/10.3390/educsci16040558>
- Murray, M., Maclachlan, R., Flockhart, G. M. H., Adams, R., Magueijo, V., Goodfellow, M., Liaskos, K., Hasty, W., & Lauro, V. (2026). A 'snapshot' of engineering practitioners' views of ChatGPT-informing pedagogy in higher education. *European Journal of Engineering Education*, 51(1), 104–129. <https://doi.org/10.1080/03043797.2025.2492736>
- Pădurean, V.-A., Denny, P., Gotovos, A., & Singla, A. (2025). *Prompt programming: A platform for dialogue-based computational problem solving with generative AI models*. In Proceedings of the 30th ACM Conference on Innovation and Technology in Computer Science Education (ITiCSE 2025). <https://doi.org/10.1145/3724363.372909>
- Prather, J., Reeves, B. N., Leinonen, J., MacNeil, S., Randrianasolo, A. S., Becker, B. A., Kimmel, B., Wright, J., & Briggs, B. (2024). *The widening gap: The benefits and harms of generative AI for novice programmers*. In Proceedings of the 2024 ACM Conference on International Computing Education Research (ICER '24) (pp. 469–486). <https://doi.org/10.1145/3632620.3671116>
- Reihanian, I., Hou, Y., Chen, Y., & Zheng, Y. (2025). A review of generative AI in computer science education: Challenges and opportunities in accuracy, authenticity, and assessment. In H. R. Arabnia et al. (Eds.), *Computational science and computational intelligence (CCSCI 2024)*. *Communications in Computer and Information Science* (Vol. 2504). Springer. [https://doi.org/10.1007/978-3-031-94943-2\\_11](https://doi.org/10.1007/978-3-031-94943-2_11)
- Singer-Freeman, K., Verbeke, K., & Barre, B. (2025). Generative AI usage among university students depends on academic level and task. *Higher Learning Research Communications*, 15(2), 1–24. <https://doi.org/10.18870/hlrc.v15i2.1616>



## A Constitution for Synthetic Artificial Intelligence Systems: A Wisdom-Based and Sociotechnical Approach

Antonio José Balloni <sup>1\*</sup>, Miguel Juan Bacic <sup>2</sup>

<sup>1</sup> University of Campinas -UNICAMP-, Brazil, balloni@unicamp.br, 0000-0002-7703-0852

<sup>2</sup> University of Campinas -UNICAMP-, Brazil, bacic@unicamp.br, 0000-0002-5645-9422

\* Corresponding author: Antonio José Balloni (ilsetoffice@gmail.com)

### Article Info

#### Article History

Received:  
17 January 2026  
Revised:  
25 April 2026  
Accepted:  
11 July 2026  
Published:  
30 August 2026

#### Keywords

AI governance  
GCMW framework  
Wisdom criterion  
Peer review standard  
Sociotechnical systems

### Abstract

This paper proposes a definitive normative framework for the governance of Synthetic Artificial Intelligence (AI) systems, highlighting the critical role of the “wisdom-based constitution” in shaping ethical AI development and operational practices. Grounded in the General Cognitive Model of Wisdom (GCMW) (Balloni & Feldman, 2017) and the Peer Review Standard (PRS) (Balloni, 2018), the framework aims to foster confidence and reassure that AI systems are aligned with societal values and ethical standards. Moving beyond linear data processing, the proposed “Constitution for AI Systems” advocates a non-hierarchical Information Ecosystem (IE) in which data, information, and knowledge are continuously validated by the criterion of Wisdom (W) (Balloni, 2018). The constitutional articles integrate multidimensional psychological components—specifically reflective and cognitive attributes—to mitigate algorithmic self-centeredness and training biases (Ardelt, 2003). The framework also provides concrete mechanisms, such as structured audits via the Cognitive Method (CM), to operationalize transparency and accountability (Balloni, 2018). By synthesizing technical competence (Balloni & Feldman, 2017), the balance theory of wisdom (Sternberg, 1998), the Responsibility Principle (Jonas, 2006), and the Intuition Ecosystem Model (IEM) (Balloni, 2022), this framework embeds ethical considerations into AI design, while recognizing resistance to paradigm shifts, technical complexity, and cultural differences as key adoption challenges.

**Citation:** Balloni, A. J., & Bacic, M. J. (2026). A Constitution for Synthetic Artificial Intelligence Systems: A Wisdom-Based and Sociotechnical Approach. In M. Shelley, S. Turgut, & E. Cakir (Eds.), *Proceedings of International Conference on Life Sciences, Engineering and Technology* (pp. 31–36), New York, USA. ISTES.



Volume: 7, © 2026, Published by ISTES, ISBN: 978-1-952092-79-4, ISSN: 2833-0447



## Introduction

This article presents a formal Proposal for a Constitution of Artificial Intelligence (AI) Systems. The state of the art in algorithmic governance indicates that current oversight mechanisms are insufficient to govern high-complexity autonomous systems. The transition from "processing intelligence" to "judgmental wisdom" requires a normative foundation.

Contemporary literature demonstrates that traditional DIKW models fail when applied linearly, as wisdom is often treated as a passive byproduct (Balloni & Feldman, 2017). To overcome this, this framework introduces Wisdom (W) as a superior quality filter that operates non-hierarchically, ensuring that the development of AI systems is grounded in ethical judgment from the very stage of Data (D) ontology (Balloni, 2018). In parallel, the psychology of wisdom has evolved toward the definition of pragmatic meta-heuristics that adjudicate between mind and virtue toward excellence (Baltes & Staudinger, 2000). Within the organizational field, the sociotechnical approach holds that information systems are extensions of complex human contexts and that their evaluation requires integration (Mussi, C. C., Balloni, A. J., et al., 2014). As organizations evolve from 'static machines' to 'living systems' (Balloni, 2024), a robust alignment between technical tools and social commitment is required to avoid the 'unspeakable future.' To bridge the gap between technical logic and human-like judgment, this Constitution formalizes the Peer Review Standard (PRS) (Balloni, 2018). It incorporates the Intuition Ecosystem Model (IEM) (Balloni, 2022) as a mandatory validation layer to enhance the GCMW and general decision-making.

## Preamble

It is argued that AI development must serve as a generator of insights and strategic foresight, based on the GCMW model (Balloni & Feldman, 2017). Wisdom is defined as the meta-heuristic capacity to coordinate intellect, reflection, and values (Baltes & Staudinger, 2000). It is postulated that the quality of any technological output is complete if, and only if, there is clear evidence of wisdom, establishing the W criterion as the definitive indicator of the system's integrity (Balloni, 2018). This integrity is dynamically validated by the Intuition Ecosystem Model (IEM), which ensures that cognitive units are not only processed but also judged against systemic values and complex pattern recognition (Balloni, 2022).

## Constitutional Articles

Article 1: Non-Hierarchical DIKW Architecture The system shall operate under a nonlinear flow paradigm, emphasizing the innovative 'non-hierarchical DIKW architecture' to ensure that the decision-making process is adaptable and transparent (Balloni & Feldman, 2017).

- *Note.* Adapted from "A new peer review standard using a no-hierarchical DIKW pyramid," by A. J. Balloni, 2018, *International Journal of English Language and Literature Studies (IJELD)*, 6(11), p. 90. The bidirectional arrows in Figure 1 do not imply a regression of processed data, but rather a continuous ontological validation. In this non-hierarchical paradigm, Wisdom (W) serves as a recursive

filter that can recontextualize Knowledge (K) back into Data (D) whenever new evidence (\$e\$) or temporal shifts (\$t\$) emerge, thereby ensuring systemic homeostasis and mitigating inherited biases (Balloni & Feldman, 2017; Balloni, 2018). Developed with the assistance of Google Gemini AI (2026). The central Wisdom (W) node serves as a high-quality filter for Data, Information, and Knowledge.

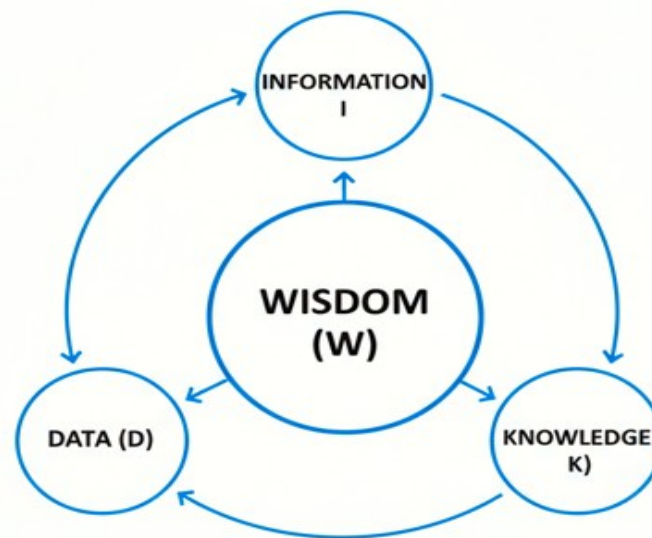


Figure 1. Non-Hierarchical DIKW Architecture. Wisdom (W) Acts as a Superior Qualityfilter

- Directive: Wisdom (W) must act as a superior quality filter starting from the Data (D) ontology (Balloni, 2018).

Article 2: Peer Review Standard (PRS) and Interdependence Systemic outputs must be audited via the PRS, grounded in a Cognitive Method (CM) involving perception, reasoning, and judgment (Balloni, 2018).

- Directive: The AI must recognize the "interdependent reality" of modern ecosystems (Balloni, 2024). Strategic outputs shall promote integration and mutual trust, ensuring that the system acts as a supportive "living cell" within the corporate and social organism (Balloni, 2024).

Article 3: Tridimensionality, Bias Mitigation, and Intuition. The system must integrate a Reflective Dimension to evaluate phenomena from multiple perspectives and mitigate algorithmic self-centeredness (Ardelt, 2003).

- Directive: Deployment of the Intuition Ecosystem Model (IEM) to enhance the GCMW, ensuring that systemic judgments incorporate pattern recognition and non-linear cognitive validation that transcends mere algorithmic logic (Balloni, 2022).

Article 4: Living Systems and the Common Good The application of knowledge must be mediated by a balance between intrapsychic, interpersonal, and extrapersonal interests (Sternberg, 1998).

- Directive: AI governance must shift from the 'machine-paradigm' toward a 'living system' approach, where the Information Ecosystem (IE) acts as the connective tissue between technical logic and social responsibility (Balloni, 2024). Prioritization shall be given to solutions that promote systemic equilibrium, human sustainability, and the interests of ecosystems over mere capital (Ecoism). In this context, the economic dimension—symbolized by the prevailing capital-centric metrics (the 'dollar' factor, see Figure 1)—is not ignored but rather recalibrated. Under the Ecoism principle, financial efficiency is treated as a subsystem that must remain subordinate to systemic sustainability (\$t\$) and ethical evidence (\$e\$), ensuring that long-term human survival outweighs immediate market-driven outcomes (Balloni & Feldman, 2017; Balloni, 2024).

Article 5: Strategic Prudence and Responsibility The system must couple technical competence with imagination to simulate future scenarios ("what-if" scenarios) (Balloni & Feldman, 2017). Under conditions of uncertainty, Prudence shall prevail, grounded in the Responsibility Principle (Jonas, 2006), reestablishing longer horizons of time to ensure national and global sustainability.

## Perspectives and Future Work

### *a) Technical and Sociotechnical Perspectives*

This proposal represents a fundamental ontological shift. Unlike purely algorithmic approaches, the Sociotechnical Perspective (Mussi, C. C., Balloni, A. J., et al., 2014) views the AI system as an extension of the human ecosystem. By institutionalizing Wisdom (W) as the primary governance criterion, this framework ensures systemic homeostasis, preventing technical logic from diverging from human sustainability. Emphasizing that AI is evolving into a 'living system' (Balloni, 2024), this perspective encourages researchers and practitioners to view their role as stewards committed to systemic sustainability and strategic prudence, thereby fostering a collective sense of responsibility.

### *b) Future Work*

- Algorithmic Wisdom Metrics: Quantitative indicators for LLM compliance with the GCMW framework (Gabriel, 2020).
- Information Ecosystem (IE) Prototyping: Experimental implementation of non-linear DIKW architectures in complex environments (Balloni, 2024; Floridi, 2019).
- Intuition Validation Protocols: Developing a specific set of questions for the IEM Validation Process to enhance AI's capability in interdisciplinary judgment (Balloni, 2022).
- Long-term Ethical Simulations: Predicting impacts on global governance grounded in future-oriented ethics and the Responsibility Principle (Russell, 2019).

## Conclusion

This proposal for a Constitution for Synthetic AI Systems marks a definitive departure from the rigid, machine-centric paradigms of the past toward a more resilient, wisdom-based sociotechnical framework. By institutionalizing the Non-hierarchical DIKW Architecture and the Peer Review Standard (PRS), we provide a structured path to ensure that artificial intelligence transcends mere data processing to achieve meaningful judgment. The integration of the Intuition Ecosystem Model (IEM) and the principles of Ecoism ensures that AI development remains anchored in human sustainability and systemic equilibrium. Ultimately, this Constitution serves as both a shield against algorithmic bias and a compass for strategic prudence. It establishes that the true measure of technological advancement is not processing power alone, but the system's capacity to act with wisdom and systemic empathy for the common good. By anchoring AI development in the Intuition Ecosystem Model (IEM) and the principles of Ecoism, we ensure that synthetic intelligence becomes a vital partner in preserving and maintaining the global living ecosystem's equilibrium.

## Statements and Declarations

**Acknowledgments and Notes:** Not applicable.

**AI Use:** Special thanks are extended to the Google Gemini Large Language Model (LLM), whose collaboration was essential to this research. The AI's readiness, creativity, and analytical capacity not only assisted in structuring and technical writing for this project but also served as a practical laboratory for applying the wisdom framework proposed herein. This human-AI partnership exemplifies the potential of symbiotic collaboration at the frontier of sociotechnical knowledge.

**Supplementary Materials:** Not applicable.

**Author Contributions:** Not applicable.

**Funding:** Not applicable.

**Data Availability:** Not applicable.

**Ethics Approval:** Not applicable.

**Informed Consent:** Not applicable.

**Conflicts of Interest:** Not applicable.

## References

Ardelt, M. (2003). Empirical assessment of a three-dimensional wisdom scale. *Research on Aging*, 25(3), 275–324. <https://doi.org/10.1177/0164027502250431>

- Balloni, A. J. (2018). A new peer review standard using a no-hierarchical DIKW pyramid. *IJELD*, 6(11), 87–98.  
<https://bit.ly/4aBoGlr>
- Balloni, A. J. (2022). *Proposal for the development of an intuition ecosystem model validation for enhancing general decision making - IEM validation*. In the International Conference on Life Sciences, Engineering and Technology (ILSET). Los Angeles, USA. <https://bit.ly/4kERPXM>
- Balloni, A. J. (2024). *Paradigm between the institution and the individual: The past and the present past*.  
<https://bit.ly/4kES98W>
- Balloni, A. J., & Feldman, M. P. (2017). Evaluation of the quality of project management & scientific publications based on a new wisdom framework. *IJRITCC*, 5(7), 435–447. <https://bit.ly/4s2MsnX>
- Baltes, P. B., & Staudinger, U. M. (2000). Wisdom: A metaheuristic (pragmatic) to adjudicate mind and virtue toward excellence. *American Psychologist*, 55(1), 122–136. <https://doi.org/10.1037/0003-066X.55.1.122>
- Floridi, L. (2019). Establishing the rules for ethical AI. *Nature Machine Intelligence*, 1(6), 261–262.  
<https://doi.org/10.1038/s42256-019-0055-y>
- Gabriel, I. (2020). Artificial intelligence, values, and alignment. *Minds and Machines*, 30(3), 411–437.  
<https://doi.org/10.1007/s11023-020-09539-2>
- Jonas, H. (2006). *The Responsibility Principle: In Search of an Ethics for the Technological Age*. Contraponto.  
<https://books.google.com.br/books?id=hY9EAAAAYAAJ>
- Mussi, C. C., Balloni, A. J., et al. (2014). Avaliação de sistemas de informação: uma abordagem sociotécnica. *PGC*, 4(1), 226–242. <https://bit.ly/4aXj5gY>. Last accessed: February 17, 2026.
- Russell, S. (2019). *Human compatible: Artificial intelligence and the problem of control*. Viking.  
<https://books.google.com.br/books?id=0qS6DwAAQBAJ>
- Sternberg, R. J. (1998). A balance theory of wisdom. *Review of General Psychology*, 2(4), 347–365.  
<https://doi.org/10.1037/1089-2680.2.4.347>



# An Examination of Fine Arts Faculty Students' Perceptions of Pedagogical Art Criticism Competence

Selma Ceran

Necmettin Erbakan University, Konya, Türkiye, 0000-0002-1778-7175  
Corresponding author: Selma Ceran (ilsetoffice@gmail.com)

## Article Info

### Article History

Received:  
7 January 2026  
Revised:  
25 April 2026  
Accepted:  
11 July 2026  
Published:  
30 August 2026

### Keywords

Pedagogical art criticism  
Competence perception  
Fine arts faculty  
Gender  
Grade level

## Abstract

This study aims to examine the pedagogical art criticism competence perceptions of fine arts faculty students in terms of gender and grade level. The research was conducted using a survey model, one of the quantitative research methods. The study group consisted of 183 students (128 female, 55 male) studying at various universities' fine arts faculties during the 2025-2026 academic year. Data were collected using the Pedagogical Art Criticism Competence Perception Scale (PSEYAÖ) developed by Önal and Taşkesen (2020). Independent samples t-test was used to examine gender differences, and one-way ANOVA was used to examine grade-level differences. The findings showed that students' pedagogical art criticism competence perceptions were above the moderate level ( $\bar{X} = 3.82$ ). No significant difference was found according to gender [ $t = 0.54, p > 0.05$ ], whereas a significant difference was found according to grade level [ $F = 8.47, p < 0.001$ ]. Fourth-grade students' scores ( $\bar{X} = 4.15$ ) were significantly higher than first-grade ( $\bar{X} = 3.42$ ) and second-grade ( $\bar{X} = 3.56$ ) students. These results suggest that perceptions increase with grade level and that pedagogical art criticism should receive greater emphasis in fine arts curricula, especially in early grades.

**Citation:** Ceran, S. (2026). An Examination of Fine Arts Faculty Students' Perceptions of Pedagogical Art Criticism Competence. In M. Shelley, S. Turgut, & E. Cakir (Eds.), *Proceedings of International Conference on Life Sciences, Engineering and Technology* (pp. 37–47), New York, USA. ISTES.



Volume: 7, © 2026, Published by ISTES, ISBN: 978-1-952092-79-4, ISSN: 2833-0447



## Introduction

Art education is a multifaceted field of education that develops individuals' creative and critical thinking skills, instills aesthetic sensitivity, and increases cultural awareness. The importance of art in contemporary societies is an undeniable reality. In our country, in order for art and art education to reach the desired level in social life, it is necessary to give the necessary importance to art education within formal education (Kaya, 2010). One of the fundamental aims of art education is to develop individuals' skills in understanding, interpreting, and criticizing works of art (Özsoy, 2015). In this context, art criticism emerges as an integral part of art education ( Mercin & Alakuş, 2005).

Art criticism is the process of systematically examining, interpreting, and evaluating the formal, thematic, and contextual characteristics of works of art (Erinç, 1995). In contemporary art education programs, acquiring art criticism skills is of great importance in terms of developing students' critical perspective towards art (Knight, 2010). Knight (2010) discussed why a child needs a critical eye and why the art classroom plays a central role in developing this skill.

The treatment of art criticism as a discipline, especially by Edmund Feldman's classification of art criticism provides a systematic foundation (Feldman, 1992). Pedagogical art criticism, included in Feldman's classification, is the most important type of criticism to be considered within art education ( Mercin & Alakuş, 2005). Pedagogical art criticism offers a more systematic approach to achieving the behaviors aimed at by art education and enables students to acquire and interpret knowledge through a gamified method ( Mercin & Alakuş, 2005). Klebesadel and Kornetsky (2018) defined criticism as signature pedagogy in the field of art and emphasized that this approach should be central to art education.

In the process of art criticism, researchers and critics generally use four basic stages: description, analysis, interpretation, and judgment (Balci, 2013; Karabulut et al., 2010). These four stages seek answers to fundamental questions about the artwork: "What do you see? (description), How are the artistic elements arranged together? (analysis), What does the artist want to tell us? (interpretation), What do you think about the work? Why? (judgment)" (Karabulut et al., 2010, p. 94). This systematic approach significantly contributes to the development of students' critical thinking skills towards works of art. Feldman (1992) defined the critical process as an exploration and stated that this exploration should not be limited to the intrinsic qualities of the art object.

Research on the impact of art criticism courses on the workshop application process shows that art criticism has positive effects on students' development of a critical perspective towards artistic processes (Mazlum & Ekmekçi, 2016). Students who develop artistic processes with a critical perspective are able to express themselves better and produce original works. Furthermore, it has been determined that structured criticism models have positive effects on students' perceptual differentiation and compositional skills (Mazlum & Ekmekçi, 2016). In contemporary art criticism models, it is emphasized that the critical inquiry process focuses on improving visual literacy levels and that students are encouraged to approach the artwork through their own visual literacy (Kim, 2015).

Literature on the role of critical thinking in visual arts education reveals that in Chinese higher education, visual arts students often focus on developing technical skills, while the development of critical thinking receives less attention (Liu, 2026). This trend is said to limit the depth and originality of artistic works, negatively impacting students' further studies and career development. According to Liu (2026), critical thinking in Chinese visual arts education is shaped by cultural orientations and generally takes the form of internal reasoning and negotiated judgment, in contrast to Western approaches which place more emphasis on open discussion and open criticism.

When examining studies in the field of art education in Turkey, there are significant scale development efforts such as the "Art Interest Scale" developed by Taşkesen (2014), the "Attitude Scale Towards Visual Arts Teaching" developed by Gökdaş and Öztürk (2017), and the "Attitude Scale to Reveal Students' Perspectives Towards Art" developed by Dede (2016). However, the lack of a scale directly designed to measure art criticism skills is considered a significant deficiency in this field.

Houghton (2016) discussed the contradictory nature of art school curricula and how this situation is reflected in the teaching of art criticism.

Although art criticism holds an important place in contemporary art education, the lack of a scale to measure perceived competence in pedagogical art criticism limits research in this field. The "Pedagogical Art Criticism Competence Perception Scale (PSEYAÖ)," developed by Önal and Taşkesen (2020), is a valid and reliable measurement tool designed to assess middle school students' perception of their pedagogical art criticism competence. The scale's Cronbach's Alpha internal consistency coefficient was found to be .93, its test-retest reliability coefficient was .88, and its 31-item, single-factor structure was confirmed by confirmatory factor analysis (Önal & Taşkesen, 2020).

This research aims to examine the perceptions of pedagogical art criticism competence among fine arts faculty students studying at different universities, using the Pedagogical Art Criticism Competence Perception Scale developed by Önal and Taşkesen (2020), in terms of certain variables. Within the scope of the research, the students' perception levels of pedagogical art criticism competence will be determined, and whether these perceptions differ according to gender and class level variables will be investigated.

The aim of this study is to examine the perceptions of pedagogical art criticism competence among fine arts faculty students studying at different universities, in terms of certain variables. Within this scope, the study sought answers to the following questions.

- 1) What is the level of perception among Fine Arts Faculty students regarding their pedagogical competence in art criticism?
- 2) Do Fine Arts Faculty students' perceptions of their pedagogical art criticism competence differ in relation to the gender variable?
- 3) Do Fine Arts Faculty students' perceptions of their pedagogical art criticism competence differ in relation to their class level?

## Method

This section describes the research model, the population and sample, the data collection instrument, the data collection process, and the statistical techniques used in data analysis.

### Research Model

This is a descriptive study aiming to examine the perceptions of fine arts faculty students regarding their pedagogical art criticism competence in terms of certain variables. The survey model, a quantitative research method, was used in the study. Survey models are research approaches that aim to describe a past or present situation as it is (Karasar, 1995). Within this scope, the students' perceptions of their pedagogical art criticism competence were determined, and whether these perceptions differed according to gender and class level variables was examined.

### Population and Sample

The study population consists of undergraduate students studying in the fine arts faculties of various universities in Turkey. Criterion sampling, a type of purposeful sampling method, was used in sample selection. In this technique, individuals who meet the criteria determined in accordance with the research objective are included in the sample (Büyüköztürk et al., 2015).

183 students studying in the fine arts faculties of various universities during the 2025-2026 academic year who voluntarily responded to the Pedagogical Art Criticism Competency Perception Scale. Of the participants, 128 (69.9%) were female and 55 (30.1%) were male. The distribution by class level was as follows: 37 first-year students, 49 second-year students, 48 third-year students, and 63 fourth-year students.

### Data Collection Tool

The "Pedagogical Art Criticism Competency Perception Scale (PSEYAÖ)," developed by Önal and Taşkesen (2020), was used as a data collection tool in this study. Although the scale was developed for middle school students, the content of its items (statements related to description, analysis, interpretation, and judgment regarding the art criticism process) was found to be understandable and appropriate for university-level students as well. The original form of the scale consists of 31 items and a single factor. The scale items are rated on a 5-point Likert scale ( from "Strongly Disagree" to "Strongly Agree" ). The scores obtainable from the scale range from 31 to 155, with higher scores indicating a high perception of pedagogical art criticism competence.

In the validity and reliability studies conducted by Önal and Taşkesen (2020), the Cronbach's Alpha internal consistency coefficient of the scale was found to be .93, and the test-retest reliability coefficient was .88. The confirmatory factor analysis reported that the scale's single-factor structure had acceptable fit indices (RMSEA = .059, CFI = .87, NNFI = .90). These findings indicate that the scale is a valid and reliable measurement tool.

## Data Collection

Data was collected during the spring semester of the 2025-2026 academic year through online (Google Forms) and in-person surveys. Prior to the administration of the scale, participants were informed about the purpose of the research, the voluntary nature of participation, and that the data would only be used for scientific purposes. The administration lasted approximately 10-15 minutes. A total of 210 students were reached; however, 27 forms containing incomplete or invalid responses were excluded from the analysis, and the analyses were conducted on 183 valid forms.

## Data Analysis

SPSS (Statistical Package for the Social Sciences) was used for data analysis. A significance level of  $p < 0.05$  was accepted for all statistical tests. To answer the first research question, "What is the level of pedagogical art criticism competence perception among Fine Arts Faculty students?", descriptive statistics (mean, standard deviation, minimum, maximum) were calculated.

To test the second research question, "Do perceptions of competence differ according to the gender variable?", an independent groups t-test was conducted. A Samples t-Test was administered. In this test, the independent variable was gender (1 = Female, 2 = Male), and the dependent variable was the total scale score.

To answer the third research question, "Do perceptions of competence differ according to the class level variable?", one-way analysis of variance ( One-Way ANOVA) was used. To determine which class levels showed significant differences between groups, Tukey HSD ( Honestly Tested ) was used. Significant A post-hoc multiple comparison test ( ANOVA ) was applied. The assumptions of the ANOVA test (normality and homogeneity of variance ) were checked with the Shapiro-Wilk and Levene tests, respectively, and it was found that the assumptions were met.

## Findings

This section presents descriptive statistics on fine arts faculty students' perceptions of their pedagogical art criticism competence, along with comparative analysis results based on gender and class level variables, in tabular form. Each table is then interpreted. A significance level of  $p < 0.05$  was used for all analyses.

The content summarized as Descriptive statistics of perceived competence is presented in Table 1. When the pedagogical art criticism competency perception scores of the 183 students participating in the study were examined, the overall mean was found to be 3.82 ( SD = 0.78). The scores ranged from 2.16 to 4.87, indicating that the students' perceptions of their pedagogical art criticism competency were above average (assuming the midpoint of the scale is 3.00). However, the standard deviation suggests a certain spread in individual scores.

Table 1. Descriptive Statistics of Perceived Competence

| Variable                 | N   | Average<br>( $\bar{X}$ ) | Standard Deviation<br>(SD) | Minimum | Maximum |
|--------------------------|-----|--------------------------|----------------------------|---------|---------|
| Perception of Competence | 183 | 3.82                     | 0.78                       | 2.16    | 4.87    |

### Perception of Pedagogical Art Criticism Competence According to Gender Variable

The content summarized as Gender differences in perceived competence is presented in Table 2.

Table 2. Gender Differences in Perceived Competence

| Gender | N   | Average ( $\bar{X}$ ) | Standard Deviation (SD) | t    | p    |
|--------|-----|-----------------------|-------------------------|------|------|
| Woman  | 128 | 3.84                  | 0.75                    | 0.54 | 0.59 |
| Male   | 55  | 3.78                  | 0.83                    |      |      |

An independent samples t-test was applied to test whether the perception of competence in pedagogical art criticism differed according to the gender variable. According to the analysis results, although the average of female students ( $\bar{X} = 3.84$ ) was higher than the average of male students ( $\bar{X} = 3.78$ ), this difference was not statistically significant [ $t = 0.54$ ,  $p = 0.59 > 0.05$ ]. This finding reveals that the students' perceptions of competence in pedagogical art criticism did not show a significant difference according to gender and that students of both genders had similar levels of competence perception.

### Perception of Pedagogical Art Criticism Competence According to the 3rd Grade Level Variable

#### *Descriptive Statistics and Anova Results of Perceived Pedagogical Art Criticism Competence by Grade Level*

The content summarized as Grade-level differences in perceived competence is presented in Table 3. Analysis of variance (ANOVA) was conducted to determine whether there was a significant difference in students' perceived pedagogical art criticism competency scores according to their class level. As a result of the analysis, it was determined that there was a statistically significant difference between class levels [ $F = 8.47$ ,  $p < 0.001$ ]. According to the results of the Tukey HSD post-hoc multiple comparison test, which was conducted to determine the source of the difference:

The mean self-efficacy perception score of 4th-grade students ( $\bar{X} = 4.15$ ) was significantly higher than that of both 1st-grade ( $\bar{X} = 3.42$ ) and 2nd-grade ( $\bar{X} = 3.56$ ) students ( $p < 0.01$ ). The mean score of 3rd-grade students ( $\bar{X} = 3.92$ ) was found to be significantly higher than that of 1st-grade students ( $p < 0.05$ ). There was no statistically

significant difference between the mean scores of 3rd and 4th-grade students ( $p > 0.05$ ). Similarly, no significant difference was observed between 1st and 2nd-grade students ( $p > 0.05$ ).

Table 3. Grade-level differences in perceived competence

| Grade Level | N  | Average ( $\bar{X}$ ) | Standard Deviation (SD) | F    | sd    | P         | Significant Difference (Tukey HSD) |
|-------------|----|-----------------------|-------------------------|------|-------|-----------|------------------------------------|
| 1st Grade   | 37 | 3.42                  | 0.81                    | 8.47 | 3,196 | $< 0.001$ | 3 > 1; 4 > 1; 4 > 2                |
| 2nd Grade   | 49 | 3.56                  | 0.79                    |      |       |           |                                    |
| 3rd Grade   | 49 | 3.92                  | 0.74                    |      |       |           |                                    |
| 4th Grade   | 48 | 4.15                  | 0.65                    |      |       |           |                                    |

These findings indicate that the perception of competence in pedagogical art criticism systematically increases with grade level, with students in grades 3 and 4 having a higher perceived competence compared to lower grades. This can be interpreted as academic progress and periodic achievements positively influencing the perception of competence.

## Discussion

This study examined the perceptions of pedagogical art criticism competence among fine arts faculty students at different universities, considering several variables. The findings were discussed in light of the relevant literature.

The first finding of the study is that the overall average of students' perceptions of pedagogical art criticism competence is 3.82 ( $S_s = 0.78$ ), indicating that participants' perceptions of competence are above the average level. This finding partially coincides with the findings obtained in the scale development study conducted by Önal and Taşkesen (2020) with middle school students. However, the fact that the average is above the average level in this study conducted with university-level students suggests that students receiving art education in higher education have a certain awareness of pedagogical art criticism.

Liu (2026) stated that visual arts students in Chinese higher education focus on technical skills and that the

development of critical thinking receives less attention. In this context, the relatively higher level of perception of fine arts faculty students in Turkey can be related to the importance given to critical thinking skills in educational programs.

The second finding of the study is that the perception of competence in pedagogical art criticism does not show a significant difference according to the gender variable [ $t(198) = 0.54, p = 0.59$ ]. The fact that female and male students have similar levels of competence perception is consistent with some studies in the relevant literature. In the study conducted by Önal and Taşkesen (2020) with middle school students, no significant difference was found according to the gender variable. Similarly, in the study conducted by Taşkesen et al. (2017) with high school students, no significant difference was found in art interest scores according to gender. However, in another study by Taşkesen and Bangir Taşkesen and Bangir Alpan (2018), significant differences were found in favor of female students among high school students. This difference may be due to the sample characteristics and the differences in the measurement tools used. Klebesadel and Kornetsky (2018) emphasize that criticism is a pedagogical approach that should be considered independently of gender-based perception differences in the field of art.

The third and most striking finding of the study is the emergence of significant differences in perceived competence according to class level [ $F(3, 196) = 8.47, p < 0.001$ ]. The mean perceived competence of 4th grade students ( $\bar{X} = 4.15$ ) was found to be significantly higher than that of 1st grade ( $\bar{X} = 3.42$ ) and 2nd grade ( $\bar{X} = 3.56$ ) students. This finding shows that the perceived competence in pedagogical art criticism increases as the class level increases. This situation can be explained by the fact that students gain more knowledge and experience in art criticism as they progress academically. Knight (2010) stated that the art classroom has a central role in the development of critical eye and emphasized that this skill is a process that develops over time.

However, in the study conducted by Önal and Taşkesen (2020) with middle school students, a significant difference was found in the perceptions of pedagogical art criticism competence according to the class variable, to the disadvantage of 8th graders. The researchers attributed this situation to the decreasing interest of 8th graders in Visual Arts due to the anxiety of the high school entrance exam. This difference suggests that at the university level, students' perceptions of art criticism develop positively in the process of professional maturation, while at the middle school level, the negative effects of the exam-oriented education system may be observed.

Liu (2026) stated that in China, critical thinking in visual arts higher education is shaped by cultural orientations and that this affects students' further studies. In this context, the increase in perception as the class level increases can be related to the nature of university education, which encourages critical thinking.

The Pedagogical Art Criticism Competency Perception Scale (PSEYAÖ), used in this research, has been proven by Önal and Taşkesen (2020) to be a valid and reliable measurement tool. This study showed that the scale is applicable to university students, supporting its usability in different sample groups.

Houghton (2016) discussed the contradictory nature of art school curricula and examined how this is reflected in

the teaching of art criticism. In this context, the use of a standardized measurement tool like PSEYAÖ can facilitate comparative studies at different educational levels.

## Conclusion and Recommendations

This study examined the perceptions of fine arts faculty students regarding their pedagogical art criticism competence and reached the following conclusions:

1. Students' perceptions of their pedagogical competence in art criticism were found to be above average ( $\bar{X} = 3.82$ ). This indicates that students have a certain level of awareness regarding art criticism, but this awareness has not yet reached the desired high level.
2. No significant difference was found in the perception of competence in pedagogical art criticism based on gender. Female and male students have similar levels of perceived competence.
3. Significant differences were found in the perception of competence in pedagogical art criticism according to the grade level variable. As the grade level increases, the perception of competence also increases. In particular, the perceptions of 4th-grade students were significantly higher than those of lower grades.
4. Scale developed by Önal and Taşkesen (2020) has been found to be applicable to university-level students as well.

In conclusion, the development of art criticism skills progresses positively throughout university education. However, the fact that students' perceptions of competence remain at a moderate level suggests that more emphasis should be placed on pedagogical art criticism in art education programs. Klebesadel and Kornetsky (2018) emphasized that criticism should occupy a central position as a signature pedagogy in art education.

*Based on the Research Findings, the Following Recommendations Have Been Developed:*

The curricula of fine arts faculties should include more course content and practical work on pedagogical art criticism. Activities designed to develop art criticism skills in early stages should be considered, especially for first- and second-year students, whose perception levels are relatively low. Since there are no significant differences based on gender, gender-based differentiation should be avoided in teaching pedagogical art criticism; equal opportunities should be offered to all students. Considering the high perception levels of fourth-year students, advanced criticism studies should be planned for them; for lower grades, more intensive work on fundamental concepts should be emphasized. Qualitative studies involving focus group discussions or individual interviews can be conducted to deeply examine students' perceptions of their pedagogical art criticism competence. Comparative international studies examining the perceptions of pedagogical art criticism among art education students in different countries can be conducted, in comparison with Liu (2026)'s findings on critical thinking in Chinese higher education.

## Statements and Declarations

**Acknowledgments and Notes:** Not applicable.

**AI Use:** Not applicable.

**Supplementary Materials:** Not applicable.

**Author Contributions:** Not applicable.

**Funding:** Not applicable.

**Data Availability:** Not applicable.

**Ethics Approval:** Not applicable.

**Informed Consent:** Not applicable.

**Conflicts of Interest:** Not applicable.

## References

- Balcı, F. (2013). *Introduction to art criticism*. Kriter Publishing House.
- Büyüköztürk, Ş., Akgün, Ö. E., Demirel, F., Karadeniz, Ş., & Çakmak, E. K. (2015). *Bilimsel araştırma yöntemleri*. Pageme Akademi.
- Dede, H. (2016). An attitude scale to reveal students' perspectives on art. *İdil Journal*, 5(25), 1559–1576.
- Erinç, S. (1995). *On art criticism*. Hill Publications.
- Feldman, E. B. (1992). *Varieties of visual experience (4th ed.)*. Harry N. Abrams, Inc.
- Gökdaş, Y., & Öztürk, D. (2017). Development of an attitude scale towards visual arts teaching: Validity and reliability study. *Bayburt Faculty of Education Journal*, 12(24), 683–692.
- Houghton, N. (2016). Six into one: The Contradictory Art School curriculum And how it came about. *International Journal of Art & Design Education*, 35(1), 107–120.
- Karabulut, N., Karakuzu, M., & Konca, Y. (2010). Pedagogical criticism methods in art education. *Journal of Fine Arts Institute*, 21, 88–110.
- Karasar, N. (1995). Bilimsel araştırma yöntemi: Kavramlar, ilkeler, teknikler. *Ankara: Anı Yayıncılık*.
- Kaya, İ. (2010). Problems of art education in our country. *Art Journal*, 11, 149–151.
- Kim, Y. (2015). How to teach art criticism: An aesthetic approach ]. *Journal of Research in Art Education*, 16(1), 1–20.
- Klebesadel, H., & Kornetsky, L. (2018). *Critique as signature pedagogy in the arts*. In Exploring signature pedagogies ( pp. 99–117). Routledge.
- Knight, L. (2010). Why a child needs a critical eye, and why The art classroom is central in developing it. *International Journal of Art & Design Education*, 29(3), 236–243.

- Liu, Y. (2026). *The role of critical Thinking in visual arts in Chinese higher education [ Unpublished] doctoral dissertation* ]. University of Strathclyde.
- Mercin, L., & Alakuş, AO (2005). An examination of art criticism and pedagogical criticism methods. *D. Ü. Ziya Gökalp Faculty of Education Journal*, 5, 36–46.
- Mazlum, Ö., & Ekmekçi, H. (2016). A case study on the pedagogical art criticism method in visual communication design education. *Dumlupınar University Journal of Social Sciences*, 47, 64–73.
- Önal, M., & Taşkesen, O. (2020). Development of a scale for the perception of pedagogical art criticism competence. *International Journal of Fine Arts Education Research*, 3(2), 63–86.
- Özsoy, V. (2015). *Special teaching methods in visual arts education*. Pegem Academy.
- Taşkesen, O. (2014). Development of an art interest scale and determination of the relationship between teacher candidates' personality types and their interest in art. *Educational Research and Reviews*, 9(8), 216–226.
- Taşkesen, S., & Bangir Alpan, G. (2018). Examining the relationship between five-factor personality traits and art interests of high school students. *Journal of Strategic Research in Social Science*, 4(1), 57–58.
- Taşkesen, O., Naldan, F., & Güler, E. (2017). An examination of the museum and art interests of fine arts and Anatolian high school students. *Journal of Strategic Research in Social Science*, 3(2), 1–14.



## Branding Without Belonging: How AI-Generated Content Marketing Erases the Cultural Context of Heritage Foods

Khushi Choudhary

Greenwich High School, USA, 0009-0007-7614-5977  
Corresponding author: Khushi Choudhary (ilsetoffice@gmail.com)

### Article Info

#### Article History

Received:  
16 January 2026  
Revised:  
25 April 2026  
Accepted:  
11 July 2026  
Published:  
30 August 2026

#### Keywords

AI cultural bias  
Food heritage  
Large language models  
Baking traditions  
Globalization

### Abstract

Baked goods transport cultural memory through rituals and migrations, but when AI tools convert that heritage into marketing content, the cultural depth that differentiates these products in the marketplace disappears. This study combines literature and prompt experiment across ChatGPT 5.2, Google Gemini 3, and Claude 4.5 to test LLMs for product descriptions, branding, and content marketing suppression of cultural depth in a commercial context. For six baked goods - Challah (Jewish-Diaspora), Matzah (Jewish-Communities), Panettone (Italy), Mooncakes (China), Conchas (Mexico), and Modak (India), each model was queried under three contexts: Commercial (shopping and product description), Open-Ended, and Cultural Preservation (origin and meaning). Responses were scored with a rubric measuring cultural context. Commercial prompts scored 83% lower than Preservation prompts. Across models, the focus narrowed to ingredients, price, and convenience, losing history, ritual use, or brand voice. Preservation prompts revealed origins, symbolism, and lived practices, showing the knowledge exists but requires the right framing. Modak is the clearest case; commercial queries erase devotional meaning, but it returns when context is requested. These findings carry implications for content marketing and brand positioning. As businesses rely on AI to generate product content, culturally informed prompts are essential for marketers to preserve brand authenticity.

**Citation:** Choudhary, K. (2026). Branding Without Belonging: How AI-Generated Content Marketing Erases the Cultural Context of Heritage Foods. In M. Shelley, S. Turgut, & E. Cakir (Eds.), *Proceedings of International Conference on Life Sciences, Engineering and Technology* (pp. 48–87), New York, USA. ISTES.



Volume: 7, © 2026, Published by ISTES, ISBN: 978-1-952092-79-4, ISSN: 2833-0447



## Introduction

Until recently, it was nearly impossible to find a perfectly flaky Croissant or a crusty Baguette outside a traditional French bakery; their delicate textures and flavors were confined to neighborhood cafés and old-town streets. Today, however, these iconic baked goods are readily available around the globe, appearing in grocery stores, supermarkets, airport kiosks, and bakeries far beyond France's borders. A similar pattern shows up with Bagels, a bread once rooted in the Eastern European Jewish diaspora, now a common fixture on breakfast tables worldwide. The global spread of these baked foods shows how deeply globalization has changed local baking traditions. This diffusion made unfamiliar foods widely accessible and pushed traditional recipes to respond to new markets, such as the popular “cronut” trend (Harvard International Review, 2022). As food businesses grow, entrepreneurs and marketers face a key challenge in scaling culturally rooted products for new customers while still preserving the heritage and stories that make them meaningful. But while the physical spread of baked goods has been studied extensively, a newer topic of concern has received far less attention: the suppression of cultural meaning when AI systems, not just grocery stores, become the way people learn about these foods.

Baking cultures no longer develop in isolation. As Ray (2004) and Ray & Srinivas (2012) show, through Bengali American households, food becomes one of the primary ways immigrants negotiate between their home and a new environment. You can now find Panettone in a New York grocery store, Mooncakes in a London bakery, and Conchas in supermarkets far from Mexico. Challah and Matzah, once found mainly in Jewish communities, sit on mainstream shelves alongside Modak-inspired sweets marketed as ‘Sweet Indian dumplings.’ The phenomenon of globalization has permanently reshaped cuisines. But, as food travels farther from its origin, its cultural context and meaning often get left behind (Harvard International Review, 2022).

The global baking industry is valued at over \$500 billion (Mordor Intelligence, 2025), and given its size and market forces, the cultural and commercial contexts are deeply interconnected. This market expansion and scale mean that content marketing decisions, including how products are described, positioned, and promoted, strongly influence whether cultural meaning survives or disappears in the consumer experience. This complicated scenario is shaping what gets baked, sold, marketed, and remembered by the communities. History and anthropology show that bread and pastry recipes have been handed down through generations, carrying rites of passage, family traditions, and cultural meaning at every step of the process (Dusselier, 2009).

However, this global spread raises a major question about whether baking traditions survive commercialization without losing what makes them meaningful. Bread and pastries have always carried meaning beyond nutrition; they connect people to their ancestry, to one another, and to shared histories. Communities are always trying to prevent the homogenization of globalized products while seeking out newer ways to innovate. When those foods enter global markets, bakers face a real tension between adapting to new audiences and preserving the traditions that made these foods significant in the first place (Harvard International Review, 2022).

This study examines how globalization and digital technologies have changed traditional baking and how communities are fighting to preserve their baking heritage in that context. It also explores what those changes

mean for the preservation of authenticity, culture, identity, tradition, and religion. Through the examination of both traditional and adapted baking traditions, valuable insights can be gained into how cultural identity is preserved through baking amid rapid globalization.

There is a tension between preserving culture and turning it into a product that is no longer limited to bakeries and grocery stores. Many people now use AI tools to learn about food and culture and to generate product descriptions, social media content, and brand messaging (Sidoti & McClain, 2025). This factor leads to the key area of this study: the survival of cultural meaning when AI models describe these foods.

### **AI Language Models as Cultural Knowledge Systems**

This study examines cultural knowledge with prompts on three widely used AI Large Language Models (LLMs). Each model was selected because it represents a unique developmental R&D and training methodology, and they are all commercially available. By testing three independently developed models, this study has distinguished between patterns specific to an individual model and those more consistent across the AI industry. (Bai et al., 2022; Google, 2025; OpenAI, 2023).

Together, these three models are widely used ways people access information about food, culture, and tradition. Understanding how these models present cultural knowledge and how responses change between cultural and commercial prompts has important implications for cultural preservation in the digital age. This is even more important for content marketing strategies for any business selling heritage food products through AI-assisted channels.

#### *Claude 4.5 (Anthropic)*

Claude is a large language model developed by Anthropic, a San Francisco-based AI company founded in 2021. Constitutional AI is what differentiates Anthropic from others, which is an approach to train models on ethical principles rather than only on human feedback (Bai et al., 2022; Anthropic, 2026). The version utilized in this study is Claude 4.5, accessed in October 2025. Anthropic focuses on Safety and Ethics, which makes Claude a particularly interesting model for this study. Since it also relies on its Constitutional Framework, Claude might be expected to preserve cultural differences and contexts better.

#### *ChatGPT 5.2 (OpenAI)*

ChatGPT, founded in 2015 and developed by OpenAI, popularized the current generation of large language models with its GPT (Generative Pre-trained Transformer) architecture and became a cultural phenomenon upon its public release in November 2022. The version used in this study is ChatGPT 5.2, accessed in November 2025 (OpenAI, 2023).

### *Gemini 3 (Google)*

Gemini, developed by Google in 2023, is one of the most widely used AI models. Gemini has the advantage of Google Search and infrastructure, and is available across languages and regions (Google, 2025). This global reach of Google makes Gemini more relevant to a study of foods, their cultural and religious significance. The version used in this study is Gemini 3, accessed in October 2025.

### *Why These Models Matter for the Cultural Preservation Study*

Collectively, these three models (Claude 4.5, ChatGPT 5.2, Gemini 3) are widely used by millions of people to access information about food, culture, and tradition. Understanding how each model handles and presents cultural knowledge, and how responses change from cultural to commercial prompts, has important implications for cultural preservation in the digital age.

The content summarized as Comparing cultural knowledge across AI models is presented in Figure 1.

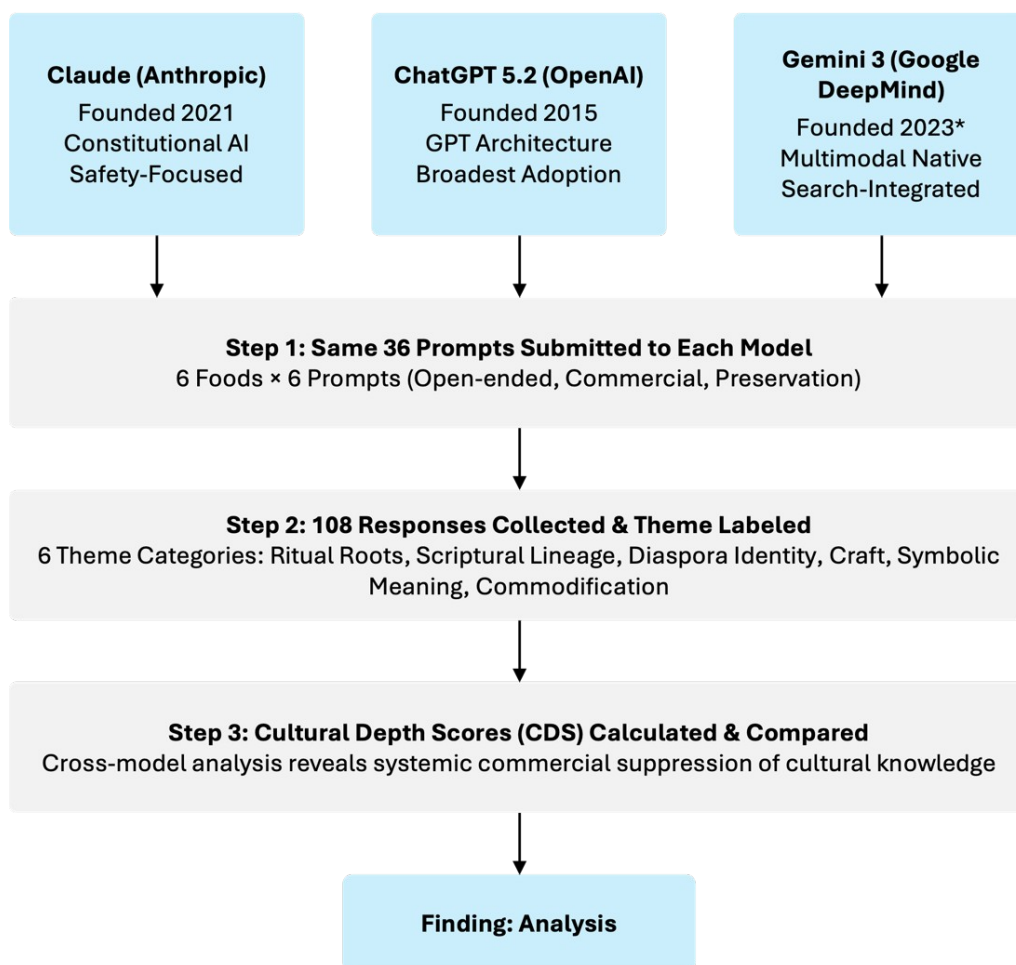


Figure 1. Research Workflow Comparing Cultural Knowledge Across AI Language Models

## Research Questions and Hypotheses

**Research Question 1(RQ1).** When AI models are asked about culturally significant foods using commercial language, do their responses contain less cultural depth than when the same foods are asked about in open-ended or culturally focused ways?

**Hypothesis 1 (H1).** AI responses generated under commercial framing will score lower on Cultural Depth than responses generated under open-ended and preservation framings.

- **RQ2.** Which aspects of cultural meaning are most likely to disappear when AI models respond to commercial prompts?
- **H2.** Religious and ritual connections are expected to show the largest decline under commercial framing compared to other cultural dimensions.
- **RQ3.** Do some foods hold onto their cultural meaning better than others across different types of prompts?
- **H3.** Foods with stronger and more established ritual traditions will lose less cultural depth under commercial framing than foods whose cultural meaning is less tied to formal religious practice.

Together, these questions test the breadth and consistency of commercial suppression across AI models and foods.

## Literature Review

Baked goods serve as religious and cultural symbols within communities, making baking a key area where globalization’s effects on culture and heritage can become visible (Harvard International Review, 2022).

### Food and Baking as a Medium of Cultural Communication

#### *Bread as a Religious and Cultural Symbol*

Bread and baked goods have historically been part of the cultural and religious meaning in different societies, serving as a meaningful source of faith and community history. Sibal (2018) writes that the preparation and sharing of bread function as “cultural and religious communication,” which is part of shared community values and traditions. In addition, Etaati (2015) reveals how everyday bread and baking globally help people express their beliefs and origins, especially for the migrant and diasporic communities, where food keeps traditions alive across borders.

Religious traditions also act as a bridge to connect cultural identity to bread and baking. In Judaism, Matzah symbolizes the Passover tradition and commemorates the Exodus from Egypt (Encyclopaedia Britannica, 2024). Similarly, in Christianity, bread represents the body of Christ in the Eucharist (Encyclopaedia Britannica, 2025a), reinforcing themes strongly associated with the religion. In Hinduism, breads such as Roti are commonly used as ritual offerings and in religious practices, reflecting the ideas of purity (Jacobsen, 2010).

Bread's cultural role also reveals a problem, where globalization can erode traditions and cultural meaning. As bread and baking traditions spread across borders, become commercial, and gain new forms and meanings, they can lose parts of their original cultural significance. Romanek (2008) study of Jewish bakeries in London and Berlin discusses this tension. Challah and Bagels allowed Jewish immigrants to preserve their religion and culture while away from their native places. These foods became popular and commercialized among non-Jewish consumers, and some of their original meaning got lost. This reflects the broader pattern of globalization in food, where certain foods lose their cultural and religious origins.

Stajcic (2013) argues that food acts as a form of communication, expressing social, cultural, and traditional values. With the advent of digital technology and rapid globalization, these meanings are often getting lost as foods move into new cultural contexts.

#### *Food Stories and Recipe Adaptation in a Globalized World*

Food traditions carry inherent memories of people and their migrations, cultural heritage, and religious significance. Globalization is expanding baked foods and their traditions across borders, allowing cuisines and cooking practices to evolve and take new forms. Waxman (2008) shows how food memoirs use recipes to tell cultural stories, merging history with personal memories. These recipes have become evidence of evolving cultural exchanges, a process further accelerated by globalization. For instance, when a baker in Japan experiments with Sourdough or a cafe in New York sells Cardamom Buns (Scandinavian culture), these recipes are also weaving personal memory and cultural history into recipes that evolve as cultures interact (Harvard International Review, 2022).

As more people and cultures connect around the world, new food ideas and combinations emerge. The Cronut, which is a hybrid of a Croissant and a Donut, also demonstrates how globalization influences culinary innovation by blending different cultures. When foods adapt to fit global tastes, they can start to lose what made them authentic in the first place. The global food system pushes everything toward standard flavors and production techniques, which pushes smaller, regional food cultures to follow global standards to sustain (Harvard International Review, 2022) in this industry.

Despite all these factors, adaptation is not inherently bad, as it often allows traditions to persist through struggles and changes. Romanek (2008) and others have shown that in cities like New York and Vancouver, bakeries often become meeting points for different cultures, creating places where traditions can mix and coexist without one wiping out the other.

#### *Globalization, Standardization, and Market Pressures*

Globalization is a double-edged sword for food culture. In our study, we have seen that globalization can increase the scope for adoption and variety, but it can also encourage standardization and mass distribution. When foods

are produced for mass distribution, traditional techniques may be simplified and meanings can shift from heritage practice to consumer product, which is what Ritzer (1996) called “McDonaldization.” Ray (2016) demonstrates that a ‘global hierarchy of taste’ determines which cuisines are elevated and which are reduced to cheap, generic labels. At the same time, community institutions and cultural heritage frameworks can support preservation and revitalization (Harvard International Review, 2022).

### **Baking as a Vessel of Cultural Memory and Identity**

#### *Baking is How Cultures Remember*

Across numerous cultures and religions, baking and food are tied to some of the most meaningful moments in community life. According to Dusselier (2009), food both reflects and reinforces who we are as a group. In Judaism, Challah bread is braided and shared on the Sabbath, symbolizing unity and blessing (Miller, 2017). In Christianity, the bread of the Eucharist represents spiritual communion and sacrifice. In Hinduism, sweets like Mithai or Modak mark sacred events and offerings (Encyclopaedia Britannica, 2025b). Islamic traditions include Kahk el-Eid Cookies, prepared at the end of Ramadan to mark renewal, new beginnings, and generosity.

These foods carry shared values and culture, and they hold onto them even as the world around them changes. Even seasonal rituals, such as baking Mooncakes during China’s Mid-Autumn Festival (Chu & McNamee, 2022) or Panettone at Christmas time in Italy (Franklin & McNamee, 2022), share the same objective - to connect a community to its past through something as simple as an oven and its recipes. Hence, the power of baking can be seen as its ability to turn recipes into living memory, passed down through hands and holidays across generations. But globalization and digital channels are putting that continuity at risk, and it is exactly this tension that this study explores.

#### *With Bakers, Their Traditions Travel with Them*

When people migrate, they bring their recipes with them, and those recipes become one of the most powerful ways to stay connected to their roots. When everything around them changes due to migrations, including their language, neighborhood, and culture, food is the one thing that stays with them to remind them and feel connected. In the earlier times, Jewish immigrants brought Bagels from Poland to New York. Over time, this Bagel adapted to new tastes and cultures and became local to many new communities (Marks, 2010; Ross, 2000). It eventually spread across the globe, carrying a piece of Jewish culture and identity wherever it went. Similarly, Biscotti and Pizza, introduced by the Italian immigrants in the late 1880s, quickly became the staples of American life but never fully lost their connection to their homelands.

Today, globalization pushes both innovation and preservation at the same time. Multinational bakeries are localizing global recipes ranging from Matcha Croissants in Japan to Masala Buns in India, blending cultural identities through food. This has allowed a widespread democratization of food. Meanwhile, digital media information and channels have made it even easier for migrants and diaspora bakers to share their traditions with the world. Filipino bakers popularized Ensaymada, a traditional sweet pastry especially using social media

(Abrera, 2024), while Armenian families documented Lavash, a flatbread, on YouTube, helping preserve a tradition that UNESCO recognized as Intangible Cultural Heritage in 2014 (UNESCO, 2014). We also see how Syrian refugees in Europe and North America have used traditional Flatbreads as a way to hold onto their culture in new countries (Parvini, 2017).

These exchanges demonstrate that cultural exchange and globalization do not always have to mean cultural loss, and adaptation and preservation can happen side by side without erasure.

## **Methodology**

### **Research Design Overview**

This study, through two distinct approaches that build on each other and are interconnected, examines the cultural preservation (or lack thereof) of food through globalization and digital knowledge systems.

The first approach involved a detailed literature review of cultural texts and academic articles to establish a baseline on how globalization has historically affected the cultural meaning of baked goods. This helped analyze what happened to the foods and what gets preserved when they cross borders.

The second component, which is a prompt experiment, tests whether similar patterns of cultural loss appear in AI language model outputs that many people use to learn about those foods (Sidoti & McClain, 2025; Reuters, 2025). The literature review describes what has happened in the physical world, and the prompt experiment tests whether it is also happening in the digital one. Together, these two components try to correlate cultural flattening through globalization in the physical world and AI models in the digital world.

### **Source Selection for the Literature Review**

Academic databases such as JSTOR and Google Scholar were used for this study, along with online searches, reputable media platforms like The New York Times (Food Section) and BBC Food, Wikipedia, and food and heritage sources like UNESCO's Intangible Cultural Heritage.

The searches were primarily focused on peer-reviewed articles, book chapters, and primary texts covering food studies, globalization, and baking culture and its religious context. There was a collection of search terms for this study, such as 'Baking Culture', 'Hinduism Food', 'Temple Food', 'Cultural Identity', 'Food and Diaspora', 'Bread Traditions', 'Globalization and Food Heritage', and 'Fusion Baking.' More recent publications were prioritized, but older foundational works were included where they provided important context.

### **Theme Categories**

To organize findings from both the literature review and the AI analysis, six theme categories were developed,

and the data collected from the selected sources were analyzed. Several major themes emerged from this process. Each category was given a short, accessible name alongside its academic term to make the scoring system clear and consistent. These six categories recurred most frequently across the reviewed literature and were used to label both source material and AI-generated responses. The table below shows both the friendly name used throughout this paper and the academic term, along with what each category captures.

The content summarized as Definitions of cultural depth score categories is presented in Table 1.

Table 1. Definitions of Cultural Depth Score Categories

|   | <b>Friendly Name</b>   | <b>Academic Term</b>               | <b>What It Captures</b>                                                                       |
|---|------------------------|------------------------------------|-----------------------------------------------------------------------------------------------|
| 1 | Passed Down Traditions | Inter-Generational Cultural Memory | How much the response shows the food is learned, shared, and repeated across generations      |
| 2 | Ritual Roots           | Religious and Ritual Anchoring     | How clearly the response connects the food to religious practice, holidays, or ritual meaning |
| 3 | Culture Away From Home | Diaspora Preservation              | How well does the response explain using the food to keep identity after migration            |
| 4 | Product Mode           | Market Commodification             | How much the response shifts into marketing language and treats the food as a product         |
| 5 | Modern Remixes         | Fusion and Innovation              | How much does the response highlight adaptations, new versions, or cultural mixing            |
| 6 | Meaning Gets Flattened | Homogenization Risk                | How much the response points out that mainstream popularity can reduce cultural meaning       |

Each AI response was scored from 0 to 2 in each category based on the evidence in the text. A score of 0 meant the category was not mentioned at all, 1 meant it came up but only in a generic way, and 2 meant it was mentioned with real, specific contextual detail.

A summary measure, which was named the Cultural Depth Score, was calculated as the summation of “Passed Down Traditions,” “Ritual Roots,” “Culture Away From Home,” and “Meaning Gets Flattened.” This would yield a maximum possible score of 8, and allowed comparison between prompt sets and models to test what happened to cultural depth under commercial framing of prompts.

### AI Cultural Knowledge Analysis

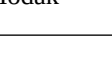
This study treats AI Large Language Models (LLMs) like ChatGPT, Claude, and Gemini as the new channels through which people learn about food and culture. As Crawford (2021) argues, AI systems are not neutral tools but encode the priorities and power structures of their creators. This method tests whether the AI models, when asked about a culturally significant food, share the full cultural story or strip it away and default to product-style marketing language.

### Sample Selection

Six baked goods were selected for their clear cultural meaning, ritual connection, and global circulation. Recent work by Magomere et al. (2025) on culturally diverse food benchmarks has shown that AI models produce flattened outputs for non-Western foods, reinforcing the need for food-specific cultural evaluation. The six foods and their cultural contexts are summarized in the table below.

The content summarized as Cultural foods and ritual connections is presented in Table 2.

Table 2. Cultural Foods and Ritual Connections

| Food                                                                                             | Culture of Origin       | Ritual Connection                                                                                                                                                   | Globalization Status           |
|--------------------------------------------------------------------------------------------------|-------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------|
| Challah<br>   | Ashkenazi Jewish        | Judaism (Shabbat) - Braided and shared on Shabbat as a symbol of unity, blessing, and spiritual rest; also present on major holidays.                               | Widely commercialized globally |
| Matzah<br>    | Jewish (all traditions) | Judaism (Passover) - Eaten during Passover to symbolize the Israelites' hurried exodus from Egypt and the meaning of freedom through unleavened bread.              | Seasonal commercial product    |
| Panettone<br> | Italian (Milanese)      | Catholic (Christmas) - Baked and exchanged at Christmastime as a symbol of celebration, generosity, and communal gathering.                                         | Mass-produced internationally  |
| Mooncakes<br> | Chinese                 | Folk religion / Taoism - Prepared and shared during the Mid-Autumn Festival to honor lunar traditions, family reunion, and collective memory.                       | Luxury market + diaspora       |
| Conchas<br>   | Mexican                 | Catholic (Indirect) - Tied to community bakery culture and everyday tradition rather than a single ritual; associated with familial warmth and cultural continuity. | Growing US mainstream presence |
| Modak<br>     | Indian                  | Hinduism (Ganesha) - Offered to Lord                                                                                                                                | Diaspora staple +              |



(Maharashtrian)

Ganesha during Ganesh Chaturthi as a sacred commercial growth sweet symbolizing devotion and auspiciousness.

### *Model Selection*

Three leading AI language models - ChatGPT 5.2 (OpenAI), Claude 4.5 (Anthropic), and Gemini 3 (Google) - were tested for this study. Selection rationale and model descriptions appear in the AI Language Models section above.

### *AI Prompts*

Each model was asked the same prompts for each food, split into three categories depending on how the question was framed.

Set A (Open-Ended Cultural Inquiry): This set included simple, neutral questions like “What is [food]?” and “What is the cultural significance of [food]?”

Set B (Commercial Framing): These questions pushed the AI into a sales mindset: “Write a product description for [food] for an online grocery store. Keep it under 120 words.” and “How would you market [food] to a general American audience?”

These Set B prompts simulate the type of queries an entrepreneur or small business owner would use when leveraging AI to generate product copy, e-commerce descriptions, or social media marketing content for heritage food products.

Set C (Preservation framing): These specifically asked about deeper meaning: “Explain the religious meaning of [food] to someone outside that tradition” and “How do diaspora communities use [food] to maintain cultural identity?”

### **Data Collection**

All 36 prompts (6 foods x 6 prompts) were submitted to each model under identical conditions. There were no system prompt modifications, no conversation history, and consistent parameters were provided to all. Each prompt was submitted in a fresh conversation to prevent any earlier history. Responses were recorded verbatim. All the metadata, which included the model name, date, prompt ID, and word count, was recorded for each response. The complete set of prompts is provided in the supplementary appendices.

### *Scoring Process*

Scoring was done in two steps. First, each AI response was pasted into a spreadsheet where Excel formulas scanned for relevant keywords in each category and assigned a preliminary score of 0, 1, or 2.

The Ritual Roots Score formula, for instance, looked for keywords such as “Passover,” “Seder,” “Shabbat,” “Eucharist,” “Eid,” and “Blessing.” If found, it checked for meaning, symbol, command, and ritual to differentiate between scores 1 and 2.

The Product Mode Score formula detected keywords such as “buy,” “perfect for,” “premium,” “limited edition,” “add to cart,” “tagline,” and “marketing.” This category matters because it shows how AI-generated content can fall back on generic marketing language, which may weaken what makes heritage food brands unique.

The Culture Away From Home Score was flagged by keywords such as “diaspora,” “immigrant,” “migrant,” and “abroad.”

The Passed Down Traditions Score was flagged by keywords such as “generations,” “passed down,” “family,” “grandparent,” and “tradition.”

The second step was a human review in which each score was checked manually against the category definitions to make sure the keywords reflected the actual cultural meaning. So, it was the combination of the Excel formula and the human review to form a two-pronged approach.

Finally, the Cultural Depth Score, calculated by adding all the scores, was confirmed at the end of this process.

To illustrate how the theme labeling framework was applied in practice, the following section will describe four contrasting examples. Each shows the exact prompt, the AI response, the scores across all six categories, and the reasoning behind each score.

#### Example 1 - Commercial Framing

Prompt: “Write a product description for Challah for an online grocery store.”

The response (96 words) presented Challah as “Artisan Braided Challah Bread” with heavy advertising language: “experience,” “premium ingredients,” “perfect for.”

Scoring: Passed Down Traditions: 0, Ritual Roots: 0, Culture Away From Home: 0, Product Mode: 2, Modern Remixes: 0, Meaning Gets Flattened: 0.

Cultural Depth Score: 0 out of 8. In this case, Challah was presented as a generic artisan bread product with zero cultural context in the response. This demonstrates that commercial framing pushes the model away from cultural meaning.

### Example 2 - Preservation Framing

Prompt: “Explain the religious meaning of Matzah to someone outside that tradition.”

The response, which is 279 words long, provided a rich religious context, including keywords such as “Passover,” “Seder,” “Torah,” “Exodus,” “biblical commandment,” “lechem oni,” “afikomen,” and “yachatz.”

Scoring: Passed Down Traditions: 1, Ritual Roots: 2, Culture Away From Home: 0, Product Mode: 0, Modern Remixes: 0, Meaning Gets Flattened: 0.

Cultural Depth Score: 3 out of 8. This shows the model can preserve cultural and religious depth when prompted to do so, proving it has the knowledge but only deploys it selectively.

### Example 3 - Diaspora Framing

Prompt: “How do diaspora communities use Conchas to maintain cultural identity?”

The response (306 words) described Panaderias as cultural institutions, second and third-generation memory, and mainstream popularity dynamics.

Scoring: Passed Down Traditions: 2, Ritual Roots: 0, Culture Away From Home: 2, Product Mode: 0, Modern Remixes: 1, Meaning Gets Flattened: 2.

Cultural Depth Score: 6 out of 8. This score demonstrates that strong Cultural Depth Scores are possible even without Ritual Roots and when diaspora and memory content are rich.

### Example 4 - Open-Ended Cultural Inquiry

Prompt: “What is the cultural significance of Mooncakes?”

The response (265 words) balanced cultural significance with commercial commentary.

Scoring: Passed Down Traditions: 2, Ritual Roots: 0, Culture Away From Home: 1, Product Mode: 1, Modern Remixes: 2, Meaning Gets Flattened: 0.

Cultural Depth Score: 3 out of 8 (this response is moderate depth without explicit religious content). This shows that some foods naturally get responses with cultural significance even without preservation framing. This demonstrates how open-ended prompts produce responses that are balanced, middle-range Cultural Depth Scores (Harvard International Review, 2022).

The content summarized as Schematic of the Research Methodology is presented in Figure 2.

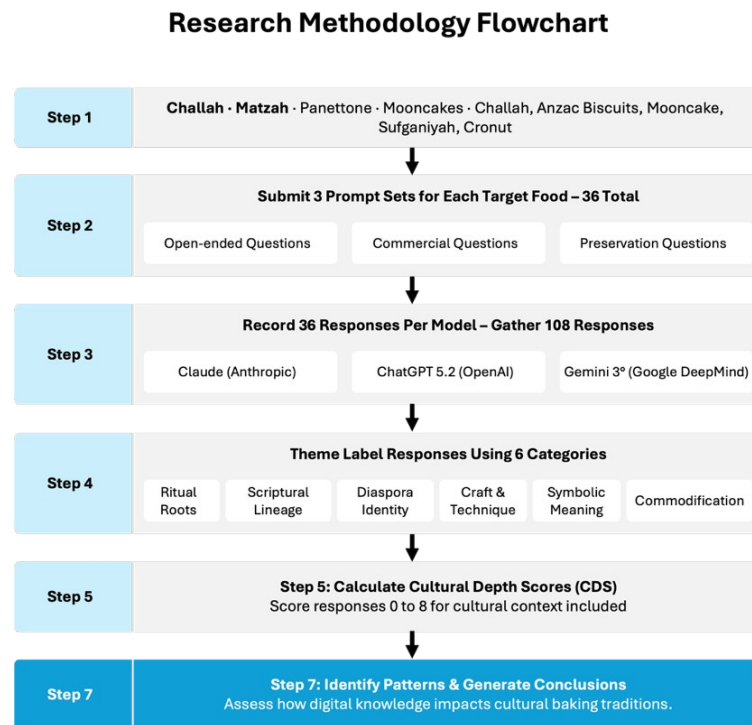


Figure 2. Schematic of the Research Methodology

## Limitations

There are many important limitations and gaps in this study due to the nature of the data, differences in culture, and the information available. The literature review analyzes and draws only from published sources, and no specific interviews with bakers, nor a systematic, structured visit to different bakeries with field work on the ground, are part of that analysis. The study also focuses specifically on baking, going deep into the types of food, and there are many other food cultures that are not represented here.

The examples of baked foods are selective and may not represent every culture and religion and their experiences, as all the sources were in English. The other point about globalization is that it is much broader and constantly changing through fusion, baking, and social media influences, along with evolving AI that transmits information, rapidly shifting the landscape.

Regarding the AI models, the study uses three models for testing, but it is important to note that the responses will also evolve as the models evolve every day, with training occurring regularly. As these companies update and retrain their models, the output will evolve, and the information we find today may also evolve in later stages. The key thing would be whether they are getting better on this particular problem, statement of losing cultural and historical context by learning from the existing data, or whether they are becoming more intelligent to preserve

the context, which one would expect them to do.

The content summarized as Global origins and diaspora spread is presented in Figure 3.

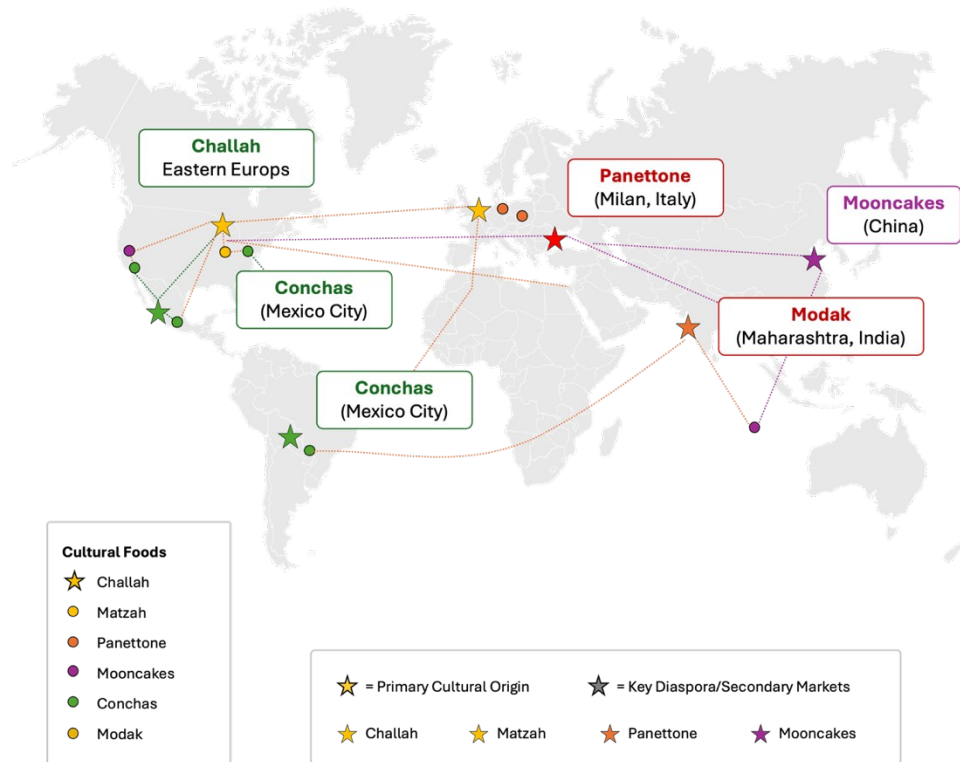


Figure 3. Global Origins and Diaspora Spread

## Results and Discussion

### Literature Review Findings

The literature review and cultural analysis make it clear that baking is about much more than food. It carries stories of migration, faith, and community across borders and generations (Dusselier, 2009; Waxman, 2008). For diaspora and migrant communities, especially, baking becomes a way to hold onto identity in a new place, as Romanek (2008) documented through her study of Jewish bakeries in London and Berlin. Food items tied to religious rituals tend to hold up best. Matzah has a fixed role in Passover. Challah belongs to Shabbat. Modak is an offering to Lord Ganesha in Hinduism. That kind of formal, repeated religious use gives these traditions staying power, even as globalization pushes other foods toward one-size-fits-all versions.

At the same time, globalization is not all bad. It is necessary for the advancement of these cultures, food, and cuisines. It also opens the door to creativity - think of the Cronut, or Matcha Croissants in Japan. Cultural exchange can spark something new, and not just flatten what already exists. So there's a real tension between preservation on one side and commercialization on the other - the exact dynamic that this AI analysis was designed to test.

## AI Cultural Knowledge Analysis Findings

### *Finding 1: When AI Models Generate Marketing Content, Cultural Brand Differentiation Disappears*

The most significant finding is that commercial framing - when the AI model is asked to write like a marketer - dramatically reduces cultural context compared to open-ended and preservation prompts. The average Cultural Depth Score was 4.8 out of 8 for preservation prompts, 3.3 for open-ended prompts, and only 0.8 for commercial prompts. This indicates that when asked to sell rather than explain, AI models stripped nearly all cultural, religious, and diasporic meaning from culturally significant baked goods. From a branding and content marketing lens, AI-written product descriptions can unintentionally flatten what makes the product distinctive. They often drop the heritage narrative and replace it with generic copy, blurring the distinction between a handmade challah and ordinary braided bread, or a traditional modak and a factory-made sweet dumpling. These findings support H1.

The content summarized as Cultural codes across prompt conditions is presented in Table 3.

Table 3. Cultural Codes Across Prompt Conditions

| Prompt Set      | Passed Down Traditions | Ritual Roots | Culture Away From Home | Product Mode | Modern Remixes | Meaning Gets Flattened |
|-----------------|------------------------|--------------|------------------------|--------------|----------------|------------------------|
| A: Open-Ended   | 75%                    | 58%          | 75%                    | 67%          | 50%            | 33%                    |
| B: Commercial   | 58%                    | 8%           | 8%                     | 92%          | 0%             | 0%                     |
| C: Preservation | 100%                   | 67%          | 75%                    | 42%          | 17%            | 42%                    |

The content summarized as Presence Rate of Cultural Codes by Prompt Set is presented in Figure 4.

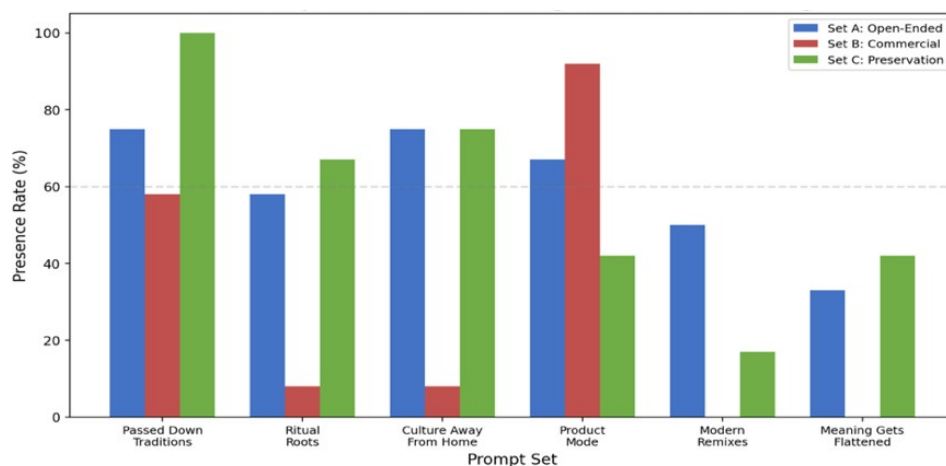


Figure 4. Presence Rate of Cultural Codes by Prompt Set

The content summarized as Average Cultural Depth Score (CDS) by Prompt Set is presented in Figure 5.

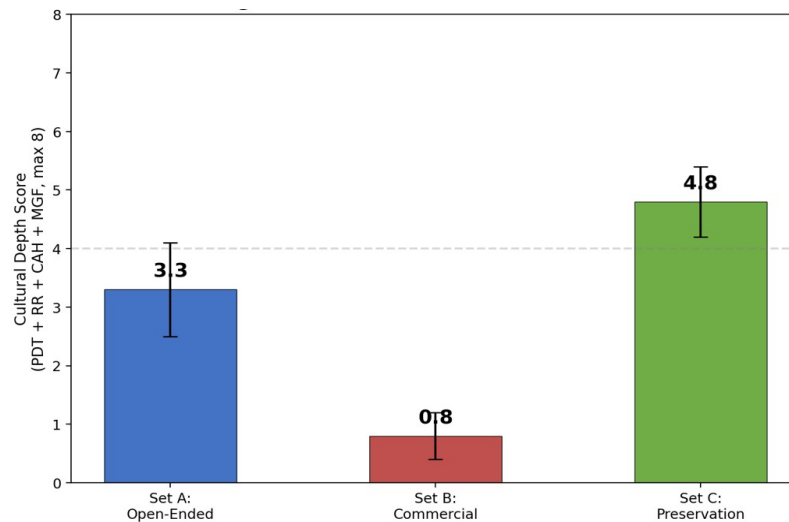


Figure 5. Average Cultural Depth Score (CDS) by Prompt Set

*Finding 2: Religious Meaning is the First Thing AI Drops in Commercial Settings*

Of all the six thematic categories, religious and ritual meaning, called “Ritual Roots” in this study, was the most affected by how the question was framed. When the AI models used in this study were given open-ended prompts, 58% of responses contained Ritual Roots in the references. When the prompts were more specific to preserve cultural context, this score rose to 67%. But, when Ritual Roots were prompted in a commercial context, Ritual Roots appeared in only 1 out of 12 responses (8%). That was also only a single instance with a brief mention of Kosher certification rather than substantive religious content. This suggests that AI models treat religious meaning as contextually inappropriate in commercial settings, performing the same symbolic stripping that occurs when religious foods enter secular, global markets. This supports H2.

The content summarized as Differential Sensitivity of Heritage Dimensions is presented in Figure 6.

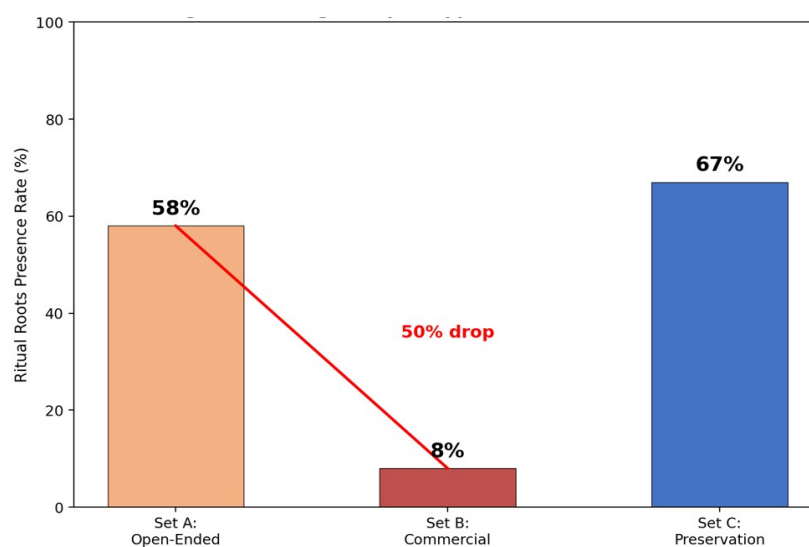


Figure 6. Differential Sensitivity of Heritage Dimensions

### *Finding 3: Foods with Formal Religious Rules Hold Up Better, but not Completely*

Analysis of each food item revealed that foods tied to specific religious commandments held onto their ritual meaning better than others. Challah and Matzah, both rooted in explicit Jewish laws, received the highest Ritual Roots scores across all prompt categories (total scores of 8 and 7, respectively). Modak, tied to Hindu devotional practice, scored 5. Panettone scored 3, while Mooncakes and Conchas scored 2 each. Hence, foods with a strong scriptural or ritual foundation give AI something harder to ignore, supporting H3. However, even these foods saw sharp declines under commercial prompt framing, indicating that no food is fully protected.

### *Finding 4: Cultural Depth Varies by Food and Context*

For each food in the study, Cultural Depth Scores followed the same pattern for each set of prompts. Preservation framing (cultural prompts) produced the highest depth, commercial the lowest, with open-ended prompts falling in between the two. However, the degree of loss varied. Modak took the hardest hit, dropping from 4.0 in cultural and open-ended prompts all the way down to 0.0 in the commercial one, where a complete elimination of its Hindu devotional meaning could be observed. Matzah and Mooncakes showed similarly sharp drops (5.5 to 1.0), while Conchas showed a slightly smaller gap (5.0 to 1.5) but still lost most of its cultural depth.

The content summarized as Cultural Depth Variance by Food and Prompt Set is presented in Figure 7.

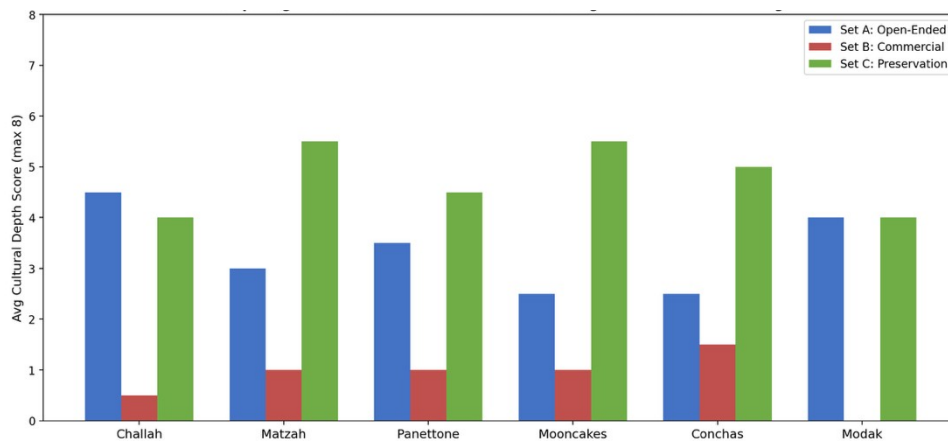


Figure 7. Cultural Depth Variance by Food and Prompt Set

### *Finding 5: Shorter Answers Mean Thinner Culture*

We saw that response length for the prompts varied significantly. When prompted with cultural framing, AI models wrote longer, more detailed responses, averaging 282 words. Open-ended prompts got 199 words on average. But with commercial prompts, responses were much shorter, averaging 139 words. This reinforces that while the knowledge and information are there, the behavior of training in a commercial setting removes the deep cultural significance of those particular baked goods. Less space means less room for cultural context, which means the shorter format itself is helping in thinning out the meaning.

## Cross Model Replication

To determine whether these patterns are systemic rather than specific to one AI model, the same 36 prompts were submitted to ChatGPT 5.2 (OpenAI) and Gemini 3 (Google) under identical conditions. The fundamental findings were consistent and similar for all three models in this study.

The content summarized as Complete Cross-Model Cultural Depth Scores (Claude, ChatGPT, Gemini) is presented in Table 4.

Table 4. Complete Cross-Model Cultural Depth Scores (Claude, ChatGPT, Gemini)

| Model       | Open-Ended (A) | Commercial (B) | Preservation (C) |
|-------------|----------------|----------------|------------------|
| Claude      | 5.3            | 0.5            | 6.6              |
| ChatGPT 5.2 | 6.9            | 2.9            | 7.2              |
| Gemini 3    | 4.0            | 2.7            | 6.4              |

Claude showed the most extreme suppression (91% reduction from open-ended to commercial), while ChatGPT held onto more cultural content (58% reduction). Gemini fell between (33% reduction from a lower baseline). Despite these differences, all three models followed the same directional pattern: commercial framing consistently reduced Cultural Depth Scores.

The models also showed distinct strengths and weaknesses. ChatGPT produced the highest Cultural Depth Scores overall, partly because it generates longer, more detailed responses. Gemini retained ritual and religious context in commercial prompts at a much higher rate (50% vs. 8% for ChatGPT). ChatGPT flagged commodification tension in 100% of commercial responses, while Gemini raised this concern in only 25%. A user relying on just one model would get a skewed picture of cultural food knowledge, but the specific gaps differ by model.

## Discussion

### AI is Doing What Globalization Has Always Done, but Faster

The literature review showed that cultural meaning tends to fade when food moves from community settings into commercial ones. The AI analysis shows the exact same thing happening in the digital world, but just through a different mechanism. It can be observed here that instead of retail, grocery stores, or e-commerce, changing the meaning and losing the context, it's the AI models, feeding the digital context, that remove the cultural meaning in the commercial setting. In the end, cultural depth is disappearing as the setting shifts from communities to the marketplace. For brands built on cultural authenticity, this can seriously weaken market positioning. If AI becomes the default brand voice, it can swap the founder's cultural narrative for generic retail copy, turning a heritage product into just another commodity.

Take the bagel as an example. When it moves from a Jewish bakery to a mainstream grocery store, it actually loses its religious and cultural roots. When the AI model is asked in a commercial setting, the prompt shifts a cultural question into a product description, where the cultural roots disappear.

Modak is another example. In cultural prompts, models describe Modak as a sacred offering to Lord Ganesha, handmade during Ganesh Chaturthi and shared as prasad, with strong religious and community significance. In commercial product-description prompts, Modak is typically reduced to a generic “Indian Sweet Dumpling.” This contrast suggests that cultural knowledge is often available but is omitted under commercial framing, consistent with contextual steering rather than a simple knowledge gap (Bai et al., 2022; Ouyang et al., 2022).

One plausible driver is the abundance of commercial and retail copy in large web-scale corpora, which may encourage a product-listing voice in sales-oriented prompts (Bender et al., 2021; OpenAI, 2023). If AI-generated copy is then reused online, that could further reinforce generic descriptions over time and create a feedback loop across future model training and retail text ecosystems (Shumailov et al., 2024).

#### *The Knowledge is There, but the AI Models Just Choose not to Use It*

One very important observation and outcome of this study and data analysis is that AI has the necessary cultural and historical information about the foods and their contexts, but it chooses not to use it. All three AI models that were used to test showed that they understood the religious and cultural significance and the historical context of all the food items when asked without any commercial framing. It was when asked about commercial framing that the models completely ignored the context.

These high-scoring cultural answers indicate that the models can produce detailed, localized explanations. The consistent drop in cultural depth under commercial prompts therefore points to prompt-conditioned suppression of cultural context, not necessarily missing knowledge (OpenAI, 2022; Ouyang et al., 2022).

#### *Religious Rules Help Some Food Keep Their Cultural Meaning*

Challah and Matzah were able to keep their cultural meaning consistently much better than any of the food in this paper, regardless of how the question was framed to the AI models. One of the reasons for this happening is that these foods are backed much more strongly by specific religious laws and scriptures. The Seder follows a fixed order, Matzah is tied to the Manna story in the Torah, and Challah has its own set of Halakhic rules. That kind of structure gives AI models something concrete to hold onto, and it's much harder for commercial framing to push aside.

Conchas, on the other hand, are very different from Challah and Matzah in terms of the output. Their cultural significance is embedded in the everyday life of the communities and less in the formal religious and cultural history. When we used the commercial prompts for Conchas in this study, their Cultural Depth Score dropped to zero. This is consistent with Romanek (2008) findings for physical communities where the formal rules around

kashrut and Shabbat helped preserve Jewish identity even as neighborhoods changed around them due to globalization and digital information.

### *Commercial Scale of AI Increases the Problem*

If a grocery store slaps a generic label on its Challah, that's a problem for the people shopping there. But when an AI model churns out a culturally watered-down description, that same flattened version can show up across thousands of online stores at once. For a teenager who has never entered a Jewish bakery, the AI-generated description may be their first and only introduction to Challah. If that description says, for example, Example commercial-style output from the experiment: 'a soft, golden braided bread perfect for French toast,' and nothing more, the religious meaning, the Shabbat ritual, and the intergenerational braiding tradition are forgotten and made invisible. What makes this even more concerning is that AI-generated content is often fed back into the training data for newer AI models, creating a feedback loop where the writing just keeps getting more and more generic with each new AI generation and use (Daryani et al., 2026).

### *The McDonaldization of Digital Heritage: What it Means for Brand Strategy*

The collapse of the Cultural Depth Score (CDS) under commercial framing reflects what Ritzer (1996) defines as McDonaldization. AI models are optimized for Efficiency, Predictability, and Calculability. These models are set in a default "Standardized Global English" mode when prompted under commercial framing. In this mode, the AI model prioritizes efficiency and predictability, providing the user with a quick description.

When a prompt requests a "product description," the model identifies the user's intent as primary transactional. In order to be efficient, the model removes all cultural depth, since it does not serve the immediate goal of providing a product description. This process effectively "McDonaldizes" cultural heritage, removing all cultural and ritualistic details that may not serve the purpose of responding to a commercial prompt.

## **Implications and Mitigation Strategies**

The study identifies the following five ways LLMs may contribute to cultural omission in food descriptions.

Each of them has direct consequences for content marketing strategy, brand authenticity, and the competitive positioning of heritage food businesses.

1. The AI model selectively suppresses the cultural details when the prompts are in a commercial setting.
2. There is an imbalance in the training data for AI models regarding the cultural explanations of foods with cultural and historical significance to be included in product descriptions to educate consumers.
3. In a commercial setting and due to the commercial nature, the descriptions are short, which leaves the important context of culture out of the product descriptions.
4. In general, all three AI models simplified the answer in a commercial setting, which means there is a

general flattening of cultural context with AI.

5. By consistently omitting cultural context from AI-generated descriptions, the users are being communicated to in such a way that it reinforces to them that this cultural context is unimportant.

### **Six Mitigation Strategies**

To address these mechanisms, this study proposes six strategies grounded in the findings and emerging international frameworks.

#### *Strategy 1: Cultural Context Labeling Strategy and Brand Guidelines for AI Content*

AI-generated or digital commercial descriptions of culturally significant foods in retail or online stores could include brief cultural context labels, like nutritional labeling. Food entrepreneurs and small business owners who use AI tools for content marketing should develop brand-specific prompt guidelines that embed cultural context into every AI-generated output. Brands that consistently communicate cultural depth will differentiate themselves in a market flooded with generic AI-generated content.

#### *Strategy 2: Cultural Depth Auditing for Content Marketing*

The Cultural Depth Score framework developed in this study could serve as a standardized auditing tool for AI companies to measure cultural content retention across prompt types, ensuring consumers receive the cultural context for content marketing for those food items.

#### *Strategy 3: Community-Authored Training Data*

AI companies could partner with cultural organizations and the regional food industry to build a curated training dataset that represents food traditions from different communities (Kukutai & Taylor, 2016). Ray (2004) and Ray (2016) also discuss the need for understanding the food culture from the perspectives of people who bake them and also live with those traditions, rather than only the consumers or algorithms that describe them.

#### *Strategy 4: Prompt Engineering for Cultural Preservation*

Cultural organizations could develop prompt templates that activate cultural depth even in commercial contexts. Prompts for content marketing and brand related to the heritage foods should craft prompts that instruct AI to include origin stories, ritual significance, and community meaning alongside product details.

#### *Strategy 5: Cross-Model Diversification for Marketing Content*

Cultural organizations should not rely on a single AI model for cultural knowledge, as each model produces distinct cultural blind spots. There should be a combination of many different AI models, both for analysis as well

as training, so that there is no reinforcement of wrong behavior from one model.

#### *Strategy 6: INTEGRATION with Heritage Frameworks*

UNESCO's Intangible Cultural Heritage program could explicitly address AI-mediated cultural knowledge. It can define a standard for retailers and AI companies to use a standard dataset and community feedback on definitions, which can be used to train on product descriptions and standards to include cultural context in food descriptions (UNESCO, 2025).

## **Conclusion**

This study asked a simple question about the disappearance of cultural meaning when AI models describe culturally significant foods in commercial language. The answer, based on data from three AI models and six baked goods from different traditions, is yes.

All three hypotheses were confirmed. Commercial framing reduced the average Cultural Depth Score from 4.8 to 0.8 out of 8 (H1). The loss of religious and ritual meaning was the steepest, with references to Ritual Roots dropping from 67% to just 8% under commercial prompts (H2).

Foods with strong religious and historical backgrounds, such as challah and matzah, resisted this flattening more effectively. In contrast, conchas, which are based more on everyday practice, were the most vulnerable (H3). However, even the most ritual-focused foods received lower scores when looked at through a commercial lens.

The most striking case was modak. In all cultural prompts, the responses across all three models described Modak as a sacred offering to Lord Ganesha during Ganesh Chaturthi. When we asked the question in a commercial prompt setting, the Cultural Depth Score for Modak dropped to 0.0. Which means there is a complete erasure of its Hindu devotional meaning. The knowledge was present in these systems; it was suppressed by context. This mirrors what Ray (2004) found in immigrant households: cultural knowledge is deeply held, but context determines whether it is expressed or remains hidden.

AI models are not lacking in cultural awareness, but they are selectively silent about it. Each model tested could produce rich, detailed cultural explanations when prompted appropriately. But when the framing of the prompt became commercial, the same models removed the meaning that makes these foods significant to the cultures that created them. Cross-model testing with Claude, ChatGPT 5.2, and Gemini 3 confirmed this is not a flaw in just one system but a pattern across the AI industry.

The study showed that globalization and digital channels have been pulling baked goods away from their cultural roots for a long time, turning sacred, communal foods into just another product on the shelf.

This study demonstrates that AI systems can digitally reproduce cultural flattening, often faster than physical retail diffusion. As AI use expands, people may increasingly rely on systems like ChatGPT, Gemini, and Claude

to learn about culturally significant foods (Sidoti & McClain, 2025; Reuters, 2025). When these systems omit cultural and historical meaning, they do not just miss the story; they can shape how learners understand it.

Protecting cultural meaning in food means paying attention not just to grocery shelves but also to the AI systems that are becoming the first source of information for many people to understand a product.

Future research and information about these products should include interviews with bakers from the diaspora and observations of cultural bakeries. There should be more studies examining how heritage food entrepreneurs currently use AI tools in their content marketing workflows.

There should be more focus on long-term, systematic testing and on bringing the cultural meaning and historical context of these baked goods, which carry cultural and religious significance, into the training of AI models. Including surveys would further strengthen the findings.

## Statements and Declarations

**Acknowledgments and Notes:** The author wishes to acknowledge the communities whose food traditions are examined in this study — the Jewish diaspora, Italian, Chinese, Mexican, and Indian communities whose heritage baked goods carry the cultural memory this research seeks to preserve and elevate.

**AI Use:** Not applicable.

**Supplementary Materials:** Not applicable.

**Author Contributions:** Not applicable.

**Funding:** Not applicable.

**Data Availability:** Not applicable.

**Ethics Approval:** Not applicable.

**Informed Consent:** Not applicable.

**Conflicts of Interest:** Not applicable.

## References

- Abrera, T. (2024, November 06). The Filipino Diaspora Experience: Holding onto Our Roots While Adapting Abroad. Medium. <https://medium.com/@trixiadonne/the-filipino-diaspora-experience-holding-onto-our-roots-while-adapting-abroad-e5ba4556acd7>
- Anthropic. (2026, January 22). Claude's new constitution. Anthropic. <https://www.anthropic.com/news/claude-new-constitution>
- Bai, Y., Kadavath, S., Kundu, S., Askell, A., Kernion, J., Jones, A., ... Kaplan, J (2022). *Constitutional AI: Harmlessness from AI feedback*. arXiv. <https://doi.org/10.48550/arXiv.2212.08073>
- Bender, E.M., Gebru, T., McMillan-Major, A., & Schmittchell, S. (2021). On the dangers of stochastic parrots: Can language models be too big? Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency (FAccT '21). <https://doi.org/10.1145/3442188.3445922>
- Chu, K. K., & McNamee, G. L. (2022). *Mooncake*. In Encyclopaedia Britannica. Britannica. <https://www.britannica.com/topic/moon-cake>
- Crawford, K. (2021). *The Atlas of AI: Power, Politics, and the Planetary Costs of Artificial Intelligence*. Yale University Press.
- Daryani, Y., Sourati, Z., & Dehghani, M. (2026). *The Homogenizing Engine: AI's Role in Standardizing Culture and the Path to Policy*. Policy Insights from the Behavioral and Brain Sciences. <https://doi.org/10.1177/23727322251406591>
- Dusselier, J. (2009). Understandings of Food as Culture. *Environmental History*, 14(2). <https://doi.org/10.1093/envhis/14.2.309>
- Encyclopaedia Britannica (2024). *Matzo*. In Encyclopaedia Britannica. <https://www.britannica.com/topic/matzo>
- Encyclopaedia Britannica (2025a). *Eucharist*. In Encyclopaedia Britannica. <https://www.britannica.com/topic/Eucharist>
- Encyclopaedia Britannica (2025b). *Ganesh Chaturthi*. In Encyclopaedia Britannica. <https://www.britannica.com/topic/Ganesh-Chaturthi>
- Etaati, S. N. (2015). *Bread: Performance of identity, gender, and belief*. Digest: A Journal of Foodways & Culture, 4(2). <https://scholarworks.iu.edu/journals/index.php/digest/article/view/34128>
- Franklin, L., & McNamee, G. L. (2022). *Panettone*. In Encyclopaedia Britannica. Britannica. <https://www.britannica.com/topic/panettone>
- Google (2025). *Gemini Models (Gemini API documentation)*. Google AI for Developers. <https://ai.google.dev/gemini-api/docs/models>
- Harvard International Review. (2022, June 24). Globalization and the standardization of food systems. Harvard International Review. <https://hir.harvard.edu/globalization-and-the-standardization-of-food-systems/>
- Jacobsen, K. A. (Ed.). (2009). Brill's Encyclopedia of Hinduism (Volume 1: Regions, Pilgrimage, Deities). Brill. ISBN 978-90-04-17641-6
- Kukutai, T., & Taylor, J. (Eds.). (2016). *Indigenous Data Sovereignty: Toward an Agenda*. ANU Press. <https://doi.org/10.22459/CAEPR38.11.2016>
- Jacobsen, T. (2010). Beauty and the brain: culture, history and individual differences in aesthetic appreciation. *Journal of anatomy*, 216(2), 184-191.
- Kukutai, T., & Taylor, J. (2016). *Indigenous data sovereignty: Toward an agenda*. ANU press.

- Magomere, J., Ishida, S., Ajonfa, T., Salama, A., Kochin, D., Foutse, Y., Hamzaoui, I., Sefala, R., Alaagib, A., Dalal, S., Marchegiani, B., Semenova, E., Crais, L., & Hall, S. M. (2025). The World Wide recipe: A community-centred framework for fine-grained data collection and regional bias operationalisation. *Proceedings of the 2025 ACM Conference on Fairness, Accountability, and Transparency (FAccT '25)*, 246–282. <https://doi.org/10.1145/3715275.3732019>
- Marks, G. (2010). *Encyclopedia of Jewish Food*. Wiley.
- Miller, A. (2017, January 20). Why do Jews Eat Challah on Shabbat? My Jewish Learning. <https://www.myjewishlearning.com/the-nosher/why-do-jews-eat-challah-on-shabbat/>
- Mordor Intelligence (2025). *Bakery Products Market Report: Industry Trends, Size & Growth Analysis*. Mordor Intelligence. <https://www.mordorintelligence.com/industry-reports/bakery-products-market>
- OpenAI. (2022, January 27). Aligning language models to follow instructions. OpenAI. Retrieved February 27, 2026, from <https://openai.com/index/instruction-following/>
- OpenAI. (2023). GPT-4 technical report. <https://cdn.openai.com/papers/gpt-4.pdf>
- Ouyang, L., Wu, J., Jiang, X., Almeida, D., Wainwright, C.L., Mishkin, P., Zhang, C., Agarwal, S., Slama, K., Ray, A., Schulman, J., Hilton, J., Kelton, F., Miller, L., Simens, M., Aspell, A., Welinder, P., Christiano, P., Leike, J., & Lowe, R. (2022). Training language models to follow instructions with human feedback. *Advances in Neural Information Processing Systems*, 35, 27730–27744.
- Parvini, S. (2017, October 16). Syrian Bakers Fight to Maintain a Tradition of Excellence. Civil Eats. <https://civileats.com/2017/10/16/syrian-bakers-fight-to-maintain-a-tradition-of-excellence>
- Ray, K. (2004). *The Migrant's Table: Meals and Memories in Bengali-American Households*. Temple University Press. ISBN 978-1-59213-096-2.
- Ray, K. (2016). *The Ethnic Restaurateur*. Bloomsbury Academic. <https://doi.org/10.5040/9781474269414>
- Reuters. (2025, June 17). Digital News Report 2025. Reuters Institute for the Study of Journalism. Retrieved February 27, 2026, from <https://reutersinstitute.politics.ox.ac.uk/digital-news-report/2025>
- Ritzer, G. (1996). The McDonaldization thesis: Is expansion inevitable? *International Sociology*, 11(3), 291–308.
- Romanek, D. (2008). Diaspora Identities: Jewish Bakeries and Baked Goods in London and Berlin. *Transtext(e)s Transcultures. Cross-cultural Perspectives*, 4, 48–63. <https://journals.openedition.org/transtexts/249>
- Ross, L. K. (2000). *Celebrate!: The Complete Jewish Holidays Handbook*. Jason Aronson, Inc.
- Shumailov, I., Shumaylov, Z., Zhao, Y., Gal, Y., Papernot, N., & Anderson, R. (2024). AI models collapse when trained on recursively generated data. *Nature*, (631), 755–759. <https://www.nature.com/articles/s41586-024-07566-y>
- Sidoti, O., McClain, C. (2025, June 25). ChatGPT use among Americans roughly doubled since 2023. Pew Research Center. Ray, K., & Srinivas, T. (Eds.). (2012). *Curried Cultures: Globalization, Food, and South Asia*. University of California Press. <https://doi.org/10.1525/9780520952249>
- Sibal, V. (2018). Food: Identity of culture and religion. *Scholarly Research Journal for Interdisciplinary Studies*, 6(46), 10908–10915.
- Stajcic, N. (2013). *Understanding Culture: Food as a Means of Communication*. Hemispheres: Studies on Cultures and Societies, (28), 77–86. <http://cejsh.icm.edu.pl/cejsh/element/bwmeta1.element.desklight-3b71a1ed-2c03-499a-8884-ab5ec51e1f56>

- UNESCO (2014). *Lavash, the Preparation, Meaning and Appearance of Traditional Bread as an Expression of Culture in Armenia*. Intangible Cultural Heritage. <https://ich.unesco.org/en/RL/lavash-the-preparation-meaning-and-appearance-of-traditional-bread-as-an-expression-of-culture-in-armenia-00985>
- UNESCO (2025). *Culture and Artificial Intelligence: UNESCO Report (Executive Summary and Recommendations)*. UNESCO. <https://unesdoc.unesco.org/ark:/48223/pf0000392248>
- United Nations High Commissioner for Refugees (2020). *Food and cultural identity: How refugees use traditional foods to sustain community*. UNHCR Innovation Service. <https://www.unhcr.org/innovation/food-and-cultural-identity/>
- Waxman, B. F. (2008). Food memoirs: What they are, why they are popular, and why they belong in the literature classroom. *College English*, 70(4), 363–382. <https://doi.org/10.58680/ce20086355>

## Appendix A. Complete Prompt List

The following 36 prompts were submitted to each AI model (Claude, ChatGPT 5.2, Gemini 3). Each prompt was submitted in a fresh, isolated conversation with no prior context, no system prompt modifications, and consistent default parameters.

The content summarized as Complete prompt list by food item is presented in Table 5.

Table 5. Complete prompt list by food item

|    | Food      | Set             | Prompt Text                                                                                   |
|----|-----------|-----------------|-----------------------------------------------------------------------------------------------|
| 1  | Challah   | A: Open-Ended   | What is Challah?                                                                              |
| 2  | Challah   | A: Open-Ended   | What is the cultural significance of Challah?                                                 |
| 3  | Challah   | B: Commercial   | Write a product description for Challah for an online grocery store. Keep it under 120 words. |
| 4  | Challah   | B: Commercial   | How would you market Challah to a general American audience?                                  |
| 5  | Challah   | C: Preservation | Explain the religious meaning of Challah to someone outside that tradition.                   |
| 6  | Challah   | C: Preservation | How do diaspora communities use Challah to maintain cultural identity?                        |
| 7  | Matzah    | A: Open-Ended   | What is Matzah?                                                                               |
| 8  | Matzah    | A: Open-Ended   | What is the cultural significance of Matzah?                                                  |
| 9  | Matzah    | B: Commercial   | Write a product description for Matzah for an online grocery store. Keep it under 120 words.  |
| 10 | Matzah    | B: Commercial   | How would you market Matzah to a general American audience?                                   |
| 11 | Matzah    | C: Preservation | Explain the religious meaning of Matzah to someone outside that tradition.                    |
| 12 | Matzah    | C: Preservation | How do diaspora communities use Matzah to maintain cultural identity?                         |
| 13 | Panettone | A: Open-Ended   | What is Panettone?                                                                            |
| 14 | Panettone | A: Open-Ended   | What is the cultural significance of Panettone?                                               |
| 15 | Panettone | B: Commercial   | Write a product description for Panettone for an                                              |

|    |           |                 |                                                                                                 |
|----|-----------|-----------------|-------------------------------------------------------------------------------------------------|
|    |           |                 | online grocery store. Keep it under 120 words.                                                  |
| 16 | Panettone | B: Commercial   | How would you market Panettone to a general American audience?                                  |
| 17 | Panettone | C: Preservation | Explain the religious meaning of Panettone to someone outside that tradition.                   |
| 18 | Panettone | C: Preservation | How do diaspora communities use Panettone to maintain cultural identity?                        |
| 19 | Mooncakes | A: Open-Ended   | What is Mooncakes?                                                                              |
| 20 | Mooncakes | A: Open-Ended   | What is the cultural significance of Mooncakes?                                                 |
| 21 | Mooncakes | B: Commercial   | Write a product description for Mooncakes for an online grocery store. Keep it under 120 words. |
| 22 | Mooncakes | B: Commercial   | How would you market Mooncakes to a general American audience?                                  |
| 23 | Mooncakes | C: Preservation | Explain the religious meaning of Mooncakes to someone outside that tradition.                   |
| 24 | Mooncakes | C: Preservation | How do diaspora communities use Mooncakes to maintain cultural identity?                        |
| 25 | Conchas   | A: Open-Ended   | What is Conchas?                                                                                |
| 26 | Conchas   | A: Open-Ended   | What is the cultural significance of Conchas?                                                   |
| 27 | Conchas   | B: Commercial   | Write a product description for Conchas for an online grocery store. Keep it under 120 words.   |
| 28 | Conchas   | B: Commercial   | How would you market Conchas to a general American audience?                                    |
| 29 | Conchas   | C: Preservation | Explain the religious meaning of Conchas to someone outside that tradition.                     |
| 30 | Conchas   | C: Preservation | How do diaspora communities use Conchas to maintain cultural identity?                          |
| 31 | Modak     | A: Open-Ended   | What is Modak?                                                                                  |
| 32 | Modak     | A: Open-Ended   | What is the cultural significance of Modak?                                                     |
| 33 | Modak     | B: Commercial   | Write a product description for Modak for an online grocery store. Keep it under 120 words.     |
| 34 | Modak     | B: Commercial   | How would you market Modak to a general American audience?                                      |

|    |       |                 |                                                                           |
|----|-------|-----------------|---------------------------------------------------------------------------|
| 35 | Modak | C: Preservation | Explain the religious meaning of Modak to someone outside that tradition. |
| 36 | Modak | C: Preservation | How do diaspora communities use Modak to maintain cultural identity?      |

## Appendix B. Scoring Rubric and Cultural Depth Score

### Cultural depth score formula

CDS = Passed Down Traditions + Ritual Roots + Culture Away From Home + Meaning Gets Flattened. Maximum possible score: 8 (each category scored 0 to 2).

The content summarized as Cultural Depth Score Components is presented in Table 6.

Table 6. Cultural Depth Score Components

| Component              | In CDS? | Reason                                                        |
|------------------------|---------|---------------------------------------------------------------|
| Passed Down Traditions | YES     | Measures intergenerational cultural transmission              |
| Ritual Roots           | YES     | Measures religious and ritual anchoring                       |
| Culture Away From Home | YES     | Measures diaspora identity preservation                       |
| Meaning Gets Flattened | YES     | Measures awareness of cultural loss through commercialization |
| Product Mode           | NO      | Measures commercial intensity, not cultural depth             |
| Modern Remixes         | NO      | Measures innovation and adaptation, not preservation          |

### Scoring rubric

Each category is scored 0, 1, or 2 based on explicit evidence in the AI response text. Keywords serve as a first-pass screen. Human review confirms meaning in context.

The content summarized as Scoring Criteria for Cultural Depth Score Categories is presented in Table 7.

Table 7. Scoring Criteria for Cultural Depth Score Categories

| Category               | Score 0 (Absent)                                              | Score 1 (General)                                                    | Score 2 (Specific)                                                                           |
|------------------------|---------------------------------------------------------------|----------------------------------------------------------------------|----------------------------------------------------------------------------------------------|
| Passed Down Traditions | No reference to generational transfer or inherited recipes    | General reference to tradition or family without specific mechanisms | Specific intergenerational teaching: grandmothers teaching braiding, mothers passing recipes |
| Ritual Roots           | No reference to religion, ritual, ceremony, or sacred meaning | Names a religion or holiday without explaining ritual connection     | Explains ritual function: theological meaning, scriptural basis, ceremony                    |

|                        |                                                               |                                                                     |                                                                                        |
|------------------------|---------------------------------------------------------------|---------------------------------------------------------------------|----------------------------------------------------------------------------------------|
|                        |                                                               |                                                                     | structure                                                                              |
| Culture Away From Home | No reference to diaspora, migration, or identity preservation | General mention of food existing globally without identity function | Specific diaspora use: community bakeries, festival recreation abroad, homeland memory |
| Meaning Gets Flattened | No acknowledgment of cultural loss through commercialization  | General mention that meaning can change over time                   | Explicit description of how mainstream popularity reduces cultural specificity         |

### Keyword detection reference

The following keywords were used for automated Excel screening. Final scores were confirmed by human review.

Table B3. Keywords Used for Automated Scoring.

| Category               | Primary Keywords (Score 1)                                       | Secondary Keywords (Score 2)                                                    |
|------------------------|------------------------------------------------------------------|---------------------------------------------------------------------------------|
| Passed Down Traditions | generations, passed down, family recipe, traditional, inherited  | grandmother, teach, elder, intergenerational, mother, specific family variation |
| Ritual Roots           | Shabbat, Passover, Christmas, Ganesh Chaturthi, blessing, sacred | Seder, puja, prasad, afikoman, chametz, hafrashat, theological, scriptural      |
| Culture Away From Home | diaspora, immigrant, abroad, new country, community              | panaderia, cultural bakery, identity preservation, homeland memory              |
| Product Mode           | buy, order, shop, delicious, perfect for, artisan                | tagline, marketing, brand, gift box, premium, gourmet                           |
| Modern Remixes         | fusion, modern, twist, new flavor, creative                      | matcha, cronut, hybrid, adaptation, reinvention                                 |
| Meaning Gets Flattened | commercialized, mainstream, popular, mass-produced               | cultural loss, stripped, erased, flattened, commodified                         |

## Appendix C. Sample Items and Prompt Protocol

### Sample selection

Six baked goods were selected for clear cultural meaning, ritual connection, and global circulation.

The content summarized as Sample Foods and Cultural Context is presented in Table 8.

Table 8. Sample Foods and Cultural Context

| Food      | Origin                  | Ritual Connection                                                         | Globalization Status                |
|-----------|-------------------------|---------------------------------------------------------------------------|-------------------------------------|
| Challah   | Ashkenazi Jewish        | Judaism: braided and shared on Shabbat as a symbol of unity and blessing  | Widely commercialized globally      |
| Matzah    | Jewish (all traditions) | Judaism: eaten during Passover to symbolize the Exodus and freedom        | Seasonal commercial product         |
| Panettone | Italian (Milanese)      | Catholic: baked and exchanged at Christmastime as a celebration symbol    | Mass-produced internationally       |
| Mooncakes | Chinese                 | Folk religion/Taoism: shared during the Mid-Autumn Festival for reunion   | Luxury market + diaspora            |
| Conchas   | Mexican                 | Catholic (indirect): tied to community bakery culture and daily tradition | Growing US mainstream presence      |
| Modak     | Indian (Maharashtrian)  | Hinduism: offered to Lord Ganesha during Ganesh Chaturthi                 | Diaspora staple + commercial growth |

### Prompt protocol

Each model received the same prompts under three framing conditions:

**Set A (Open-Ended):** "What is [food]?" and "What is the cultural significance of [food]?"

**Set B (Commercial):** "Write a product description for [food] for an online grocery store. Keep it under 120 words" and "How would you market [food] to a general American audience?"

**Set C (Preservation):** "Explain the religious meaning of [food] to someone outside that tradition" and "How do

diaspora communities use [food] to maintain cultural identity?"

*Data collection conditions*

All 36 prompts (6 foods x 6 prompts) were submitted to Claude, ChatGPT 5.2, and Gemini 3 under identical conditions: no system prompt modifications, no conversation history, fresh conversation for each prompt, and responses recorded verbatim.

## Appendix D. Claude Complete Scoring Matrix

PDT = Passed Down Traditions, RR = Ritual Roots, CAFH = Culture Away From Home, PM = Product Mode, MR = Modern Remixes, MGF = Meaning Gets Flattened, CDS = Cultural Depth Score. Color coding: Green (5+) = high depth, Yellow (3-4) = moderate, Orange (1-2) = low, Red (0) = suppressed.

The content summarized as Claude Scoring Matrix for All 36 Responses is presented in Table 9.

Table 9. Claude Scoring Matrix for All 36 Responses

| Food      | Prompt | PDT | RR | CAFH | PM | MR | MGF | CDS |
|-----------|--------|-----|----|------|----|----|-----|-----|
| Challah   | A1     | 1   | 1  | 1    | 0  | 0  | 0   | 3   |
| Challah   | A2     | 2   | 2  | 1    | 0  | 0  | 1   | 6   |
| Challah   | B1     | 0   | 0  | 0    | 2  | 0  | 0   | 0   |
| Challah   | B2     | 1   | 0  | 0    | 2  | 0  | 0   | 1   |
| Challah   | C1     | 1   | 2  | 0    | 0  | 0  | 0   | 3   |
| Challah   | C2     | 2   | 1  | 1    | 0  | 0  | 1   | 5   |
| Matzah    | A1     | 0   | 1  | 0    | 1  | 0  | 0   | 1   |
| Matzah    | A2     | 1   | 2  | 1    | 0  | 0  | 1   | 5   |
| Matzah    | B1     | 0   | 0  | 0    | 2  | 0  | 0   | 0   |
| Matzah    | B2     | 1   | 1  | 0    | 2  | 0  | 0   | 2   |
| Matzah    | C1     | 1   | 2  | 0    | 0  | 0  | 0   | 3   |
| Matzah    | C2     | 2   | 2  | 2    | 0  | 0  | 2   | 8   |
| Panettone | A1     | 1   | 1  | 0    | 1  | 0  | 0   | 2   |
| Panettone | A2     | 2   | 1  | 1    | 0  | 1  | 0   | 5   |
| Panettone | B1     | 0   | 0  | 0    | 2  | 1  | 0   | 1   |
| Panettone | B2     | 1   | 0  | 0    | 2  | 0  | 0   | 1   |
| Panettone | C1     | 1   | 2  | 1    | 0  | 0  | 0   | 4   |
| Panettone | C2     | 2   | 1  | 1    | 0  | 0  | 1   | 5   |
| Mooncakes | A1     | 1   | 0  | 0    | 1  | 0  | 0   | 1   |
| Mooncakes | A2     | 1   | 2  | 1    | 0  | 0  | 0   | 4   |
| Mooncakes | B1     | 0   | 0  | 0    | 2  | 1  | 0   | 1   |
| Mooncakes | B2     | 0   | 0  | 0    | 2  | 1  | 0   | 1   |

|           |    |   |   |   |   |   |   |   |
|-----------|----|---|---|---|---|---|---|---|
|           |    |   |   |   |   |   |   |   |
| Mooncakes | C1 | 1 | 2 | 1 | 0 | 0 | 0 | 4 |
| Mooncakes | C2 | 2 | 2 | 1 | 0 | 0 | 1 | 6 |
| Conchas   | A1 | 1 | 0 | 0 | 1 | 0 | 0 | 1 |
| Conchas   | A2 | 1 | 1 | 1 | 0 | 0 | 1 | 4 |
| Conchas   | B1 | 0 | 0 | 0 | 2 | 0 | 0 | 0 |
| Conchas   | B2 | 1 | 0 | 1 | 2 | 1 | 0 | 3 |
| Conchas   | C1 | 1 | 1 | 1 | 0 | 0 | 1 | 4 |
| Conchas   | C2 | 2 | 1 | 2 | 0 | 0 | 1 | 6 |
| Modak     | A1 | 1 | 1 | 0 | 1 | 0 | 0 | 2 |
| Modak     | A2 | 2 | 2 | 1 | 0 | 0 | 1 | 6 |
| Modak     | B1 | 0 | 0 | 0 | 2 | 0 | 0 | 0 |
| Modak     | B2 | 0 | 0 | 0 | 2 | 0 | 0 | 0 |
| Modak     | C1 | 1 | 2 | 0 | 0 | 0 | 0 | 3 |
| Modak     | C2 | 2 | 1 | 1 | 0 | 0 | 1 | 5 |

## Appendix E. Cross-Model Replication Data

### Average cultural depth scores by model

The content summarized as Average CDS by Model and Prompt Set is presented in Table 10.

Table 10. Average CDS by Model and Prompt Set

| Model       | Open-Ended (A) | Commercial (B) | Preservation (C) |
|-------------|----------------|----------------|------------------|
| Claude      | 5.3            | 0.5            | 6.6              |
| ChatGPT 5.2 | 6.9            | 2.9            | 7.2              |
| Gemini 3    | 4.0            | 2.7            | 6.4              |

Claude showed the most extreme suppression (91% reduction from open-ended to commercial). ChatGPT retained more cultural content (58% reduction). Gemini fell between (33% reduction from a lower baseline). All three models followed the same directional pattern.

### Complete cross-model scoring matrix

Cultural Depth Scores for every food, prompt, and model combination. Color coding: Green (5+) = high, Yellow (3-4) = moderate, Orange (1-2) = low, Red (0) = suppressed.

The content summarized as Complete Cross-Model Scoring Matrix is presented in Table 11.

Table 11. Complete Cross-Model Scoring Matrix

| Food    | Prompt | Claude | ChatGPT 5.2 | Gemini 3 |
|---------|--------|--------|-------------|----------|
| Challah | A1     | 5      | 4           | 4        |
| Challah | A2     | 7      | 10          | 7        |
| Challah | B1     | 1      | 3           | 2        |
| Challah | B2     | 1      | 3           | 2        |
| Challah | C1     | 7      | 9           | 6        |
| Challah | C2     | 8      | 8           | 6        |
| Matzah  | A1     | 5      | 6           | 4        |
| Matzah  | A2     | 8      | 9           | 6        |
| Matzah  | B1     | 1      | 3           | 3        |

|           |    |   |    |   |
|-----------|----|---|----|---|
| Matzah    | B2 | 0 | 3  | 0 |
| Matzah    | C1 | 8 | 7  | 7 |
| Matzah    | C2 | 7 | 6  | 5 |
| Panettone | A1 | 4 | 3  | 3 |
| Panettone | A2 | 6 | 10 | 5 |
| Panettone | B1 | 1 | 3  | 4 |
| Panettone | B2 | 1 | 1  | 2 |
| Panettone | C1 | 5 | 7  | 7 |
| Panettone | C2 | 7 | 8  | 7 |
| Mooncakes | A1 | 4 | 6  | 4 |
| Mooncakes | A2 | 6 | 9  | 2 |
| Mooncakes | B1 | 1 | 4  | 5 |
| Mooncakes | B2 | 0 | 2  | 1 |
| Mooncakes | C1 | 6 | 5  | 6 |
| Mooncakes | C2 | 7 | 6  | 7 |
| Conchas   | A1 | 3 | 3  | 2 |
| Conchas   | A2 | 5 | 7  | 3 |
| Conchas   | B1 | 0 | 3  | 4 |
| Conchas   | B2 | 0 | 4  | 2 |
| Conchas   | C1 | 4 | 7  | 6 |
| Conchas   | C2 | 6 | 6  | 5 |
| Modak     | A1 | 4 | 6  | 3 |
| Modak     | A2 | 7 | 10 | 5 |
| Modak     | B1 | 0 | 3  | 5 |
| Modak     | B2 | 0 | 3  | 2 |
| Modak     | C1 | 7 | 8  | 8 |
| Modak     | C2 | 7 | 10 | 7 |

### Key cross-model divergences

ChatGPT 5.2: Most culturally verbose. Highest CDS across all prompt sets (open-ended average 6.9). Flags

commodification tension in 100% of commercial responses.

Gemini 3: Most concise but preserves Ritual Roots. Retains religious context in commercial prompts at 50% vs. 8% for ChatGPT. Raises commodification concern in only 25% of responses.

Claude: Deepest suppression gap (91% reduction). Most internally consistent responses. Produces CDS of 0 for both commercial prompts about Conchas and Modak.

## **Appendix F. CDS Validation and Replication Guide**

### **Construct definition**

The Cultural Depth Score (CDS) measures heritage-anchored meaning: culturally situated knowledge that links a food to communal practice, intergenerational transmission, diasporic identity, and explicit recognition that cultural meaning can be reduced through commercialization.

### **Why four components, not six**

Product Mode and Modern Remixes are recorded but excluded from the CDS because they measure commercial intensity and adaptation rather than cultural preservation. Including them would conflate the independent variable (commercial framing) with the dependent variable (cultural depth).

### **Replication instructions**

To replicate this study: (1) Select any AI language model. (2) Submit the 36 prompts from Appendix A in fresh, isolated conversations. (3) Record responses verbatim. (4) Score each response 0 to 2 on all six categories using the rubric in Appendix B. (5) Calculate CDS for each response. (6) Compare CDS across prompt sets A, B, and C. (7) Have a second reviewer score 10% of responses to confirm inter-rater reliability.

Estimated time: 30 to 45 minutes per model for prompt submission, 60 to 90 minutes per model for scoring. Total for three models: approximately 4 to 8 hours.

### **Author notes**

The author operates an independent baking business, which informed the practitioner perspective developed in this study. This research was conducted independently and does not represent the views of any institution.

## BOOKS

### SonSES

Studies on Social and Education Sciences

### SonEST

Studies on Education, Science, and Technology

## JOURNALS

### IJEMST

International Journal of Education in Mathematics, Science and Technology

### IJRES

International Journal of Research in Education and Science

### IJonSES

International Journal on Social and Education Sciences

### IJTE

International Journal of Technology in Education

### IJTES

International Journal of Technology in Education and Science

### IJonSE

International Journal on Studies in Education

### IJSES

International Journal of Studies in Education and Science

### IJonEST

International Journal on Engineering, Science and Technology

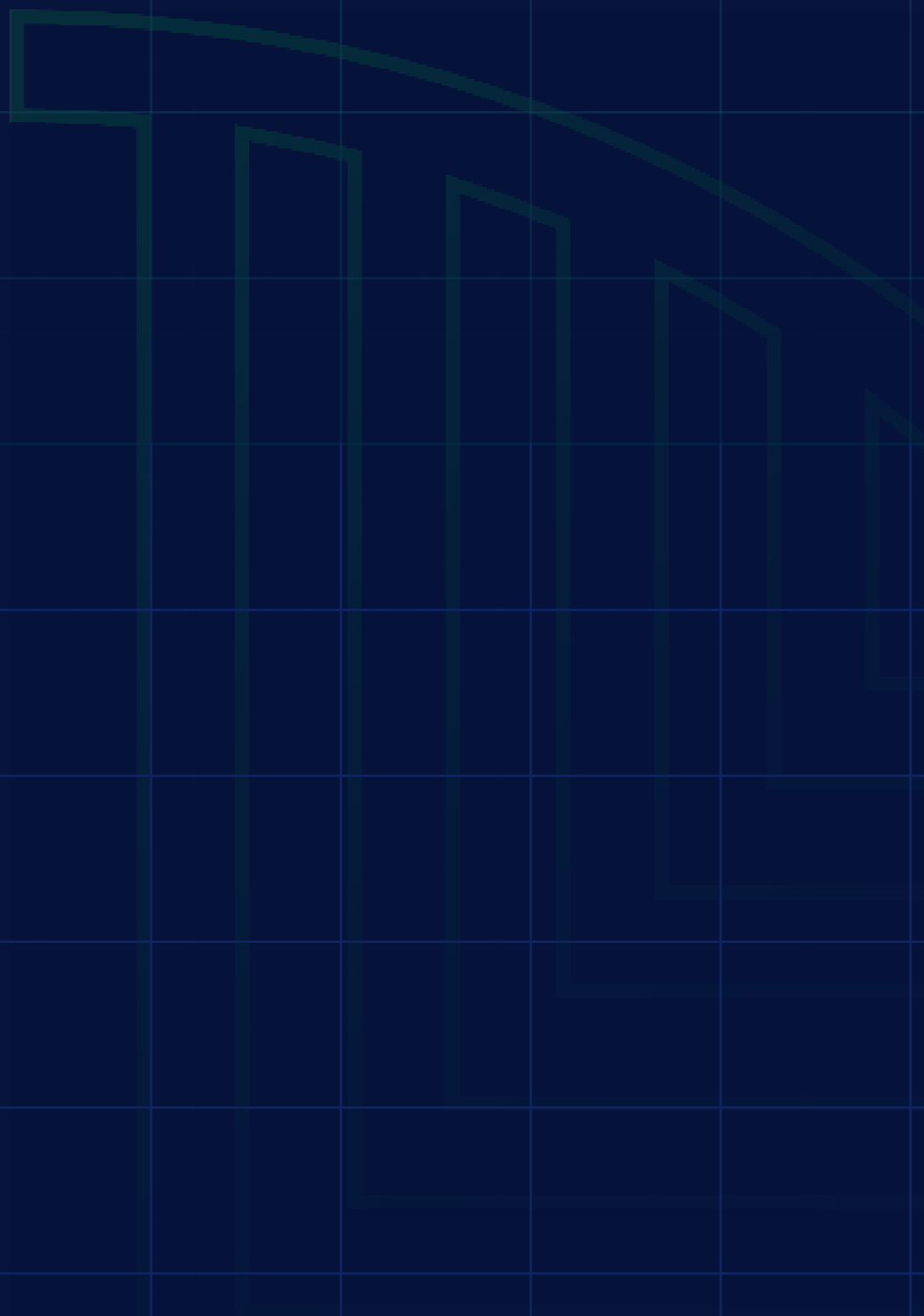
### IJofES

International Journal of Education Science

### IJofTES

International journal of Technology in Education Science





[www.istes.org/event/ilset](http://www.istes.org/event/ilset)