International Journal for Innovative Engineering and Management Research
PEER REVIEWED OPEN ACCESS INTERNATIONAL JOURNAL
www.ijiemr.org

AI-Generated Content for Digital Advertising: Transparency, Generation Failures and Human Review

Kodipaka Sairaj Goud, Indirala Abhinav, and Dudekula Teja Kala Yogeswar

Students, Department of BS-MS (Computer Science), VVISM
Hyderabad, Telangana, India | Email: sairajgoudkodipaka@gmail.com;
indiralaabhinavchary@gmail.com; tejadudekula9544@gmail.com

ABSTRACT

Businesses use AI-generated images and videos for digital advertising. Businesses use Generative AI tools to turn product photographs into advertisements. Evaluating these tools in practice requires examining job outcomes, processing time, credit deductions, and the records systems retain.

This study analysed four months of production records from one AI content-generation tool used for jewellery advertising. It covers image and video jobs, their outcomes and durations, credit deductions, and user actions such as edits, favourites, and deletions. It asks what these records show about transparency, generation failures, and human reviews.

The findings show that most failures occurred early in the period, however the records often do not say what caused them. The records show what the tool did but not who approved an output. Gaps in failure attribution, repeat-submission tracking, and structured human-review records remain. Addressing these gaps could help businesses monitor generation workflows and document decisions about AI-generated advertising, including whether outputs were reviewed before publication.

Keywords: AI image generation, digital advertising, generation failures, transparency, responsible AI, human oversight

1. INTRODUCTION

The growing adoption of AI has changed how businesses create digital advertising content. In the jewellery industry, AI tools enable marketers to convert ordinary product photographs into professionally styled images and videos. A jeweller can upload a smartphone photo of a necklace together with a reference image and obtain a finished advertisement in a short time (Figure 1). While these tools improve efficiency and reduce production costs, their use also raises important questions about transparency, reliability, and the role of human oversight in reviewing AI-generated outputs.

The tool studied here was used for jewellery advertising, so its records show how AI-generated content is produced in this industry. A jeweller who publishes these images needs answers to three simple questions. What was generated, and how much did it cost to generate? What happened when a request failed? Did anyone check the image before it went live? The last question matters because an advertisement that changes a product's stones, colour or design can mislead customers.

Jobin et al. (2019) reviewed 84 AI ethics guidelines from around the world and found that transparency and accountability were among the most common principles. Principles alone do not tell us what a real system keeps on record. This case study examines the records of one such tool.

2. REVIEW OF LITERATURE

Hartmann et al. (2025) compared AI-made and human-made advertising images in a controlled field study. An AI-generated banner got more clicks than a stock photograph, and experts rated the AI images highly. The study did not look at failed requests, charges or operational records.

Mitchell et al. (2019) proposed model cards, and Raji et al. (2020) proposed a framework for internal audits. Both are ways of documenting models and are written mainly for the people who build them. Shamsuddin et al. (2026) used Gemini to convert eight jewellery sketches into product images and asked five experts to rate them. The Mean scores ranged from 4.21 to 4.48 on a five-point scale. Their study focused on the quality of generated images, whereas the present study examines records from an operational AI content-generation service, including generation outcomes and human review activities.

3. RESEARCH GAP

The sources reviewed in this paper mainly address the quality and performance of AI-generated images, or propose ways of documenting and auditing AI models in general. Among them we found no study of the operational records of an AI tool in real use, such as job outcomes, failed requests, processing time and the credits charged for each job. Nor did we find one that asks what such records can and cannot show about transparency and human review, which matter to businesses such as jewellers that publish AI-generated images of their products. This study addresses that gap by analysing the production records of one deployed tool. The objectives that follow turn the gap into measurable questions about outcomes, timing, credits and recorded user actions.

4. OBJECTIVES OF THE STUDYs

- To examine how often AI content-generation requests are completed successfully and how often they fail.
- To investigate the time taken to generate successful outputs and determine whether recorded credit usage corresponds with the content delivered.
- To explore how generation activities and post-generation actions are recorded within the system and assess whether these records support transparency, reliability, and human oversight.

5. RESEARCH METHODOLOGY

Research Design: Retrospective descriptive case study of one AI content-generation tool.

Nature of Data: Quantitative, system-generated operational records from the tool's production database. References were used only for the literature review.

Sources of Data: A CSV export of the tool's production database with all 541 image and video job records created over four months (10 April to 25 August 2026). The export has no account identifier. Model names were taken as recorded. Some were filled in while preparing the dataset and were not checked against provider logs.

Tool Workflow: A user uploads a product photograph and a reference image to start a job. The system checks the credit balance (50 credits for an image, 150 for a video) and sends the images and a saved instruction to a hosted model. Each job's charge appears in the export as a credit amount. The user can then ask for an angle change or touch-up for 10 additional credits.

Variables and Analysis: We used job outcome and error category, assigned model, elapsed time, credits deducted and three user-action flags (edited, favourited, deleted). Elapsed time is the gap between the recorded creation and completion times, and it is available for all 463 completed jobs. For 406 jobs it comes from the recorded times. For the other 57 (12.3%), the completion time in the dataset is an estimate equal to the median recorded time for the same job type and resolution. We report counts and percentages. For elapsed time we report the median with the 10th and 90th percentiles, because a few very long times would distort an average. No statistical tests were run.

Data Use and Ethics: The export has job records only: the type of generation, its outcome, the time taken, the credits charged, and whether an output was edited, favourited or deleted. It has

no customer names or other personal details, and we did not look at any image content. We used only these job records and no other information.

Before (input)	After (output)

International Journal for Innovative Engineering and Management Research
PEER REVIEWED OPEN ACCESS INTERNATIONAL JOURNAL
www.ijiemr.org

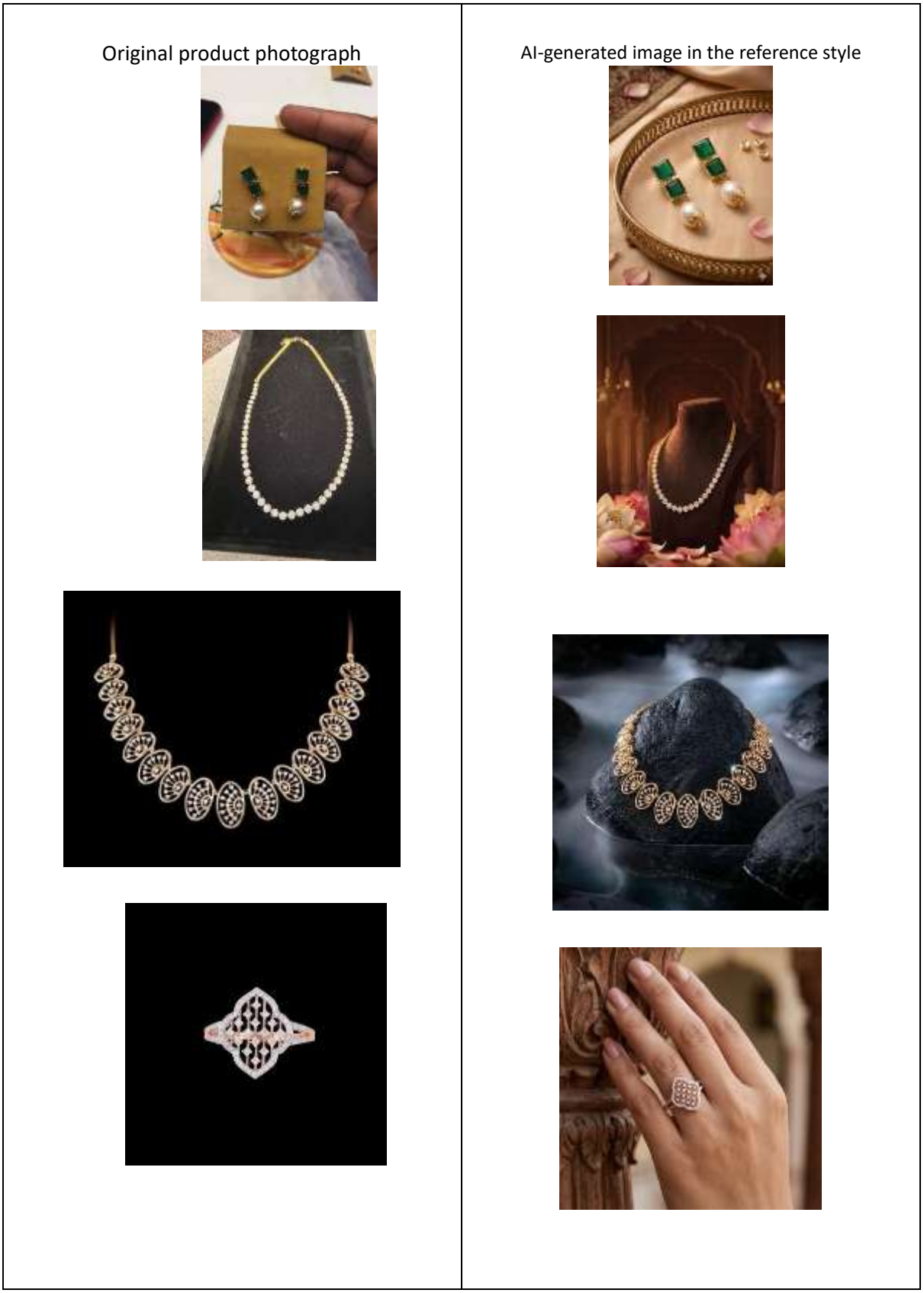


Figure 1. Example inputs (raw product photographs) and outputs generated by an AI imagegeneration model for digital marketing purposes.

International Journal for Innovative Engineering and Management Research
PEER REVIEWED OPEN ACCESS INTERNATIONAL JOURNAL
www.ijiemr.org

6. CONCEPTUAL BACKGROUND

Operational Transparency: The tool keeps enough records to trace any job, whether it succeeded or failed (Jobin et al., 2019).

Credit Transparency: Every credit charge is linked to the outcome of its job, so it can be explained.

Human-Review Records: Fields that show whether a person approved or rejected an output, such as the reviewer and the time. Without them, the database cannot show whether anyone checked an output.

7. ANALYSIS AND DISCUSSION

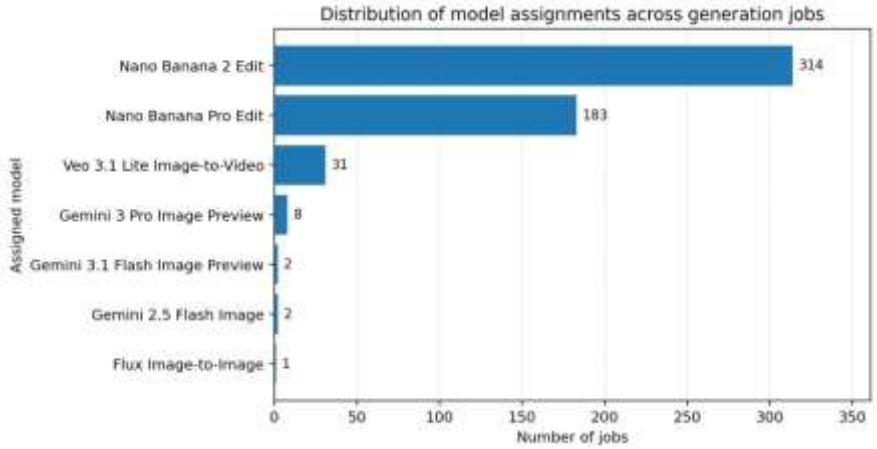


Figure 2. Distribution of model assignments across generation jobs

7.1 Job Outcomes: Of 541 jobs, 463 (85.6%) completed and 78 (14.4%) failed. Images completed more often than videos: 440 of 510 (86.3%) against 23 of 31 (74.2%). Nano Banana 2 Edit handled 314 jobs, Nano Banana Pro Edit 183, Veo 3.1 Lite Image-to-Video 31, and four other image models the remaining 13 (Figure 2).

Table 1. Elapsed time for completed jobs, including 57 estimated times
(s = seconds; pct = percentile)

Job type	n	Median (s)	10th pct (s)	90th pct (s)

Image, 2K	359	32.3	24.6	94.1
Image, 4K	81	65.9	43.7	112.2
Video, 2K	23	69.6	48.0	144.0

7.2 Elapsed Time: Table 1 gives elapsed times for all 463 completed jobs. A 4K image took about twice as long as a 2K image, and a 2K video took a little longer than a 4K image. Three 4K jobs show unrealistic times (69,905 to 238,913 s), but the median and percentiles are barely affected by them. The 57 estimated times equal the group medians, so the medians stay the same, but the spread looks slightly smaller than it really is. On the 406 recorded times alone, the 90th percentile for 2K images is 104.6 s instead of 94.1 s.

Table 2. Failed jobs by error category and month of creation, 2026

	Apr	**May**	**Jun**	**Jul**	**Aug**	**Total**
Provider validation (HTTP 422)	20	8	0	0	0	28
Webhook error, earlier version	18	8	0	0	0	26
All providers failed	5	0	0	0	0	5
Insufficient credits	0	0	2	0	0	2
Other	3	12	1	1	0	17
Failed jobs	46	28	3	1	0	78
All jobs	111	316	68	16	30	541
Failure rate	41.4%	8.9%	4.4%	6.3%	0.0%	14.4%

7.3 Generation Failures: Table 2 sorts the 78 failures by error type and month. Two types account for most of them: requests rejected by provider validation (HTTP 422) and webhook errors from an earlier version of the system, together 54 of 78 (69.2%). Both disappear after May, and only 4 failures happen from June to August.

The dataset puts 70 failures under Nano Banana Pro Edit (70 of 183 jobs, 38.3%), eight under Veo 3.1 Lite Image-to-Video (8 of 31, 25.8%) and none under Nano Banana 2 Edit (314 jobs). Much of this is due to timing. Nano Banana Pro Edit handled 179 of its 183 jobs in April and

May, when most failures happened, while Nano Banana 2 Edit was first used in May. By resolution, 62 of 444 2K jobs (14.0%) and 16 of 97 4K jobs (16.5%) failed.

7.4 Credit Records: The database holds 463 non-zero deduction records, totalling 25,450 credits. Every completed job was charged the standard price (440 images at 50 credits, 23 videos at 150), and none of the 78 failed jobs was charged. The three edited jobs show no extra charge, so any touch-up fee is not recorded in the job's own deduction.

7.5 Post-Generation User Actions: Users made 23 recorded actions: 3 edits, 4 favourites and 16 deletions, each on a different completed job. None of these is a review or approval record, and the export has no reviewer, approval, rejection or review-time fields.

7.6 Operational Transparency: The main error types suggest that problems in how the tool connects to its providers, not limits of the image models, caused most failures, although the export cannot confirm the cause. The records often say that a job failed but not which part failed it. The 17 "other" failures cannot be investigated, and no field links repeat submissions. Turnaround is fast enough for interactive use, since nine in ten timed jobs of each type finished within about two and a half minutes.

7.7 Credit Transparency: Charging followed outcomes in this export (Section 7.4), but this can be confirmed only by joining the deductions to job status, as we did, because the deduction record has no outcome field. Adding the job outcome to each deduction record would make this check easy, in line with Raji et al. (2020).

7.8 Human Oversight: Realistic images can still change a product's design, colour or stones (Figure 1). The database cannot show whether anyone checked for this. A deletion may mean the image was rejected or just tidied away, and a favourite may mean approval or simply that someone liked it. Without reviewer and decision fields, a missing review record does not prove that review was skipped. It only means the database cannot show one. Good image quality (Hartmann et al., 2025; Shamsuddin et al., 2026) and good operational records are different things, and both need to be checked.

8. MAJOR FINDINGS

- Most jobs completed, and most failures happened in the first two months.
- Provider validation errors and earlier-version webhook errors caused over two-thirds of failures, which points to problems in how the tool connects to its providers and not to model limits.

- Turnaround is fast enough for interactive use, and a 4K image takes about twice as long as a 2K image.
- Credits matched outcomes: every completed job was charged the standard price and no failed job was charged.
- The few recorded user actions are not review decisions, and the database has no reviewer fields.

9. SUGGESTIONS / RECOMMENDATIONS

- Tool designers should record model and provider IDs for every job, including failures.
- Repeat submissions should be linked with an attempt-group ID, and each deduction should store the job outcome.
- Reviewer decisions and timestamps should be added to output records.
- Organisations should check each output against the actual product before publishing it, and keep a log of job ID, reviewer, decision and date until the tool records decisions.
- Future research should combine production records with quality ratings and cover more than one tool.

10. CONCLUSION

This study looked at the production records of one AI content-generation tool. They show what the tool did well enough, and also where they stop being useful. Failure records end at an error category, credits can be matched to outcomes only by a manual join, and nothing records whether a person approved an image before it was published. Saving the outcome with each charge, linking repeat attempts and recording reviewer decisions would close these gaps and support responsible use of AI-generated advertising.

REFERENCES

1. Hartmann, J., Exner, Y., & Domdey, S. (2025). The power of generative marketing: Can generative AI create superhuman visual marketing content? *International Journal of Research in Marketing, 42*(1), 13–31.
2. Jobin, A., Ienca, M., & Vayena, E. (2019). The global landscape of AI ethics guidelines. *Nature Machine Intelligence, 1*(9), 389–399.

3. Mitchell, M., Wu, S., Zaldivar, A., Barnes, P., Vasserman, L., Hutchinson, B., Spitzer, E., Raji, I. D., & Gebru, T. (2019). Model cards for model reporting. In *Proceedings of the Conference on Fairness, Accountability, and Transparency* (pp. 220–229). ACM.
4. Raji, I. D., Smart, A., White, R. N., Mitchell, M., Gebru, T., Hutchinson, B., Smith-Loud, J., Theron, D., & Barnes, P. (2020). Closing the AI accountability gap: Defining an endto-end framework for internal algorithmic auditing. In *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency* (pp. 33–44). ACM.
5. Shamsuddin, R., Manan, E. A., Manan, M. M. A., Aziz, A. A., Jani, H. M., Muhammad, R., & Bakar, W. N. W. (2026). From jewellery design sketch to realistic product imagery using Gemini artificial intelligence. *International Journal of Research and Innovation in Social Science, 10*(1), 6821–6828.