Jurnal Ilmiah Lingua Idea
Vol 17, No 1 (2026): 132-144.
DOI: https://doi.org/10.20884/1.jli.2026.17.1.19375
ISSN 2580-1066 (online); 2086-1877 (printed)

AI as a Creative Tool in Documentary Directing "Yogyakarta Hari Ini" Developing a Prototype in the Digital Landscape Era

Diyah Verakandhi[1*], Zahrina Zatadini[2], Gregorius Arya Dhipayana[3]

[1,2,3]Institut Seni Indonesia Yogyakarta, Indonesia

verakandhi@gmail.com, zahrinazatadini@isi.ac.id, gregarya@gmail.com

Abstract: This research adopts a Practice-Led Research method by involving twenty-eight students to record fragments of urban reality in Yogyakarta from their respective perspectives. The heterogeneous collection of videos is briefly described to map the diversity of visual representations, and subsequently examined through a documentary film entitled *"Yogyakarta Hari Ini."* Specifically, this study examines how ChatGPT was employed as a narrative structuring tool: textual descriptions of the 28 footage clips were submitted to the model in randomized order, and the AI-generated sequence recommendations were subsequently evaluated and refined by the director during the editing phase. The findings demonstrate that AI effectively accelerated the categorization and narrative structuring of heterogeneous footage, while also revealing clear limitations in aesthetic judgment, atmospheric continuity, and socio-cultural sensitivity or areas that required consistent human intervention and correction. The study confirms that AI can serve as a creative tool in transforming raw visual data into a coherent narrative, functioning as a "dramaturgical assistant" during the directing phase. However, the director's aesthetic judgment remained indispensable: approximately one-third of the AI-generated sequence recommendations required revision due to mechanical transitions and disregard for atmospheric continuity. This human-machine collaborative model contributes to the discourse on Indonesian film aesthetics in the digital landscape, offering a replicable prototype for documentary practice confronting heterogeneous visual archives.

Keywords: artificial intelligence; digital landscape; documentary film; practice-led research; visual storytelling.

Submitted: 14 Jan 2026; Received in revised form: 17 June 2026; Accepted: 25 June 2026; Published regularly: 30 June 2026

To cite this article (APA Style):
Verakandhi, D., Zatadini, Z., & Dhipayana, G A. (2026). AI as a Creative Tool in Documentary Directing "Yogyakarta Hari Ini" Developing a Prototype in the Digital Landscape Era. *Jurnal Ilmiah Lingua Idea*, *17*(1) 132-144. https://doi.org/10.20884/1.jli. 2026.17.1.19375

Introduction

We exist in an era where smartphone cameras function as an extension of the human eye. Every minute, millions of moving images are recorded, circulated, archived, and consumed. This phenomenon has precipitated what is known as a visual data deluge—an abundant, almost infinitely growing volume of imagery. For documentary filmmakers, this paradigm presents a dual reality: a boon of readily accessible raw material, yet a formidable challenge in mapping reality amidst an oversaturation of data. According to industry analyses of global video traffic trends in 2024 and 2025, video content now constitutes the dominant form of data transmitted across consumer internet infrastructure worldwide, reflecting an unprecedented acceleration in moving image production that far outpaces any single director's capacity for linear review. Artificial intelligence, in this context, has emerged not as a speculative future tool but as a present-tense operational necessity in managing, interpreting, and transforming heterogeneous visual archives into coherent cinematic works.

In Indonesia, this trajectory is distinctly evident in the proliferation of community-based audiovisual productions, citizen journalism, participatory recordings, and spontaneously generated digital content. Contemporary documentary filmmakers increasingly operate not solely with self-shot footage, but with multi-source materials that are frequently non-linear, heterogeneous, and not explicitly designed for cinematic narratives. The Indonesian film industry's rapid engagement with AI tools is underscored by the emergence of events such as the Bali International AI Film Festival (BIAIFF), which in its 2025 inaugural edition showcased 44 AI-assisted films from 21 countries, including notable submissions from Indonesian creators (BIAIFF, 2025). Industry practitioners in Indonesia have reported concrete workflow transformations: VFX artists note substantial reductions in labor-intensive rotoscoping tasks previously requiring entire teams, while storyboard artists use generative AI to iterate character positions and angles in a fraction of the traditional time (Yulisman, 2025). These shifts reflect a broader national moment in which Indonesian filmmakers are actively negotiating the boundaries between human craft and algorithmic assistance—a negotiation this study directly enters.(Renov, 2004)

The *"Yogyakarta Hari Ini"*project was initiated as a collective endeavor involving 28 students within a Documentary Directing Practice course. The cohort was tasked with capturing the reality of Yogyakarta devoid of a predetermined narrative structure, guided by minimal parameters: unrestricted duration, style, and subject matter. As anticipated, the aesthetic outcome was fundamentally unconstrained.

This inherent diversity established the foundation for a methodological experiment. AI was adopted not merely as a contemporary trend, but as a practical necessity. When visual archives scale beyond the capacity of linear human review, AI offers a potential means of facilitating narrative construction during the editing phase. This study addresses two explicit research questions: (1) To what extent can ChatGPT assist in categorizing and structuring heterogeneous footage descriptions into a documentary narrative sequence? (2) Does AI function strictly as a

technical utility in documentary directing, or can it serve as a creative collaborator—and where does human intervention remain essential?

Documenter and Visual Archive in Digital Era

Bill Nichols (Nichols, 2017) defines documentary as a form of representing reality that is inherently bound to positionality. Furthermore, Verakandhi (Verakandhi, 2024) asserts that documentary film has become a medium for conveying facts and truth within an event, simultaneously capable of evoking emotion and prompting action through its ability to represent reality and construct narratives. However, in the digital era, this position is increasingly complex as documentary creation is no longer strictly director-centric.

According to Hito Steyerl (Steyerl, 2015), we are entering a period of *circulationism*, where the value of an image lies not solely in its content, but in how it circulates, is cropped, pasted, and recombined. User-generated footage amplifies this phenomenon. Crowdsourced recordings serve as evidence of public engagement while presenting distinct aesthetic challenges. Each piece of footage carries its own visual language—ranging from handheld camera shake and lens skew to spontaneous compositions and inconsistent color grading.

The evolving digital media landscape has catalyzed a shift in the role of the documentary filmmaker. While directors were once positioned as "sole observers," they now frequently operate as curators or collaborators(Nichols, 2017). We exist in a state of image surplus—abundant imagery recorded by anyone, at any time, with highly variable quality. This explosion of user-generated content (UGC) has given rise to compilation documentaries, personal archive-based films, and participatory documentaries that rely heavily on multiple perspectives.

The integration of AI into creative media workflows has attracted growing scholarly attention. Amabile and Pratt observe that computational tools are increasingly embedded in creative processes, though their role remains subordinate to human intentionaly (Amabile, T. M. & Pratt, 2016). In the context of documentary specifically, Aufderheide notes that AI-assisted editing tools are beginning to reshape post-production workflows, particularly in projects involving large and heterogeneous archives (Aufderheide, 2022). This trajectory is confirmed by recent documentary practice internationally: in Gary Hustwit's documentary Eno (2024), custom generative AI tools were used to render a different version of the film at every screening, demonstrating that AI can operate not merely as an organizational instrument but as a generative aesthetic system (International Documentary Association (IDA), 2024). Meanwhile, AI-enabled transcription and translation tools are increasingly used by documentary editors to guide assembly edits of lengthy interview footage, and directors are employing AI for audience targeting and distribution strategy. However, scholars have also raised critical concerns: AI language models such as ChatGPT operate on statistical patterns derived from training data that may embed cultural biases, potentially reflecting dominant Western aesthetic frameworks and marginalizing non-Western visual traditions(Bender, E. M.& Gebru, T.& McMillan-Major, A.& Shmitchell, 2021). Additionally, the environmental costs of large language model inference—significant energy consumption per query—represent an ethical dimension that practitioners

should acknowledge (Strubell, E. & Ganesh, A. & McCallum, 2019). Issues of bias and representation are equally salient: AI systems trained on non-diverse datasets risk amplifying stereotypical portrayals and excluding marginalized cultural perspectives (Doc Film Academy, 2026). These limitations do not disqualify AI as a creative tool, but they foreground the necessity of critical human oversight in any AI-assisted production workflow, particularly in non-Western documentary contexts where aesthetic and cultural values diverge substantially from the datasets on which most commercial AI models were trained.(Manovich, 2013)

Within this context, the primary challenge is no longer acquiring footage, but rather filtering, analyzing, and organizing these images. Directors are confronted with two increasingly distinct phases of work: first, the archival task of categorizing materials, understanding patterns, and identifying discrepancies; and second, the narrative task of determining a relevant storyline from this collection of fragments.

This research operates within this very framework. Twenty-eight students acted as a "collective recorder," generating heterogeneous visuals of Yogyakarta. This aggregation of images does not inherently yield a story; it necessitates a rigorous sorting process to curate the visuals most pertinent to the title *"Yogyakarta Hari Ini"*. Researchers must interpret visuals derived from fragmentation recorded by multiple eyes, diverse experiences, and various biases, effectively distinguishing between mere documentation and footage possessing genuine narrative potential.

Methods

Diverging from traditional arts research that remains confined to an observational standpoint, Practice-Led Research (PLR) positions creative practice as the center of gravity of the study. In this approach, the artwork is not merely a "visual example" used to illustrate a theory; rather, it becomes the intellectual space itself. The creative process transforms into a laboratory where ideas are tested, deconstructed, and reformulated through tangible action. PLR is particularly well-suited to research at the intersection of emerging technology and creative practice, because it allows the practitioner-researcher to encounter the affordances and limitations of new tools through first-hand creative engagement rather than observational distance. In the context of this study, PLR enables a direct encounter with ChatGPT as a collaborator in an actual film production—generating findings grounded in the material realities of documentary directing rather than in simulated or laboratory conditions.

PLR generates new knowledge not through critical distance, but through direct engagement. Knowledge emerges as the creator works, experiments, and subsequently reflects upon their ongoing process. This involves a cycle that is never strictly linear—iterating through experimenting, reviewing, questioning artistic decisions, and recalibrating the next steps. This is the essence of PLR: research that lives *within* the practice, rather than outside of it. (Dean, R. T. & Smith, 2009)

In this study, PLR was operationalized through four documented iterative cycles. In the first cycle, the researcher conducted manual analysis of all 28 footage clips (Phase 2), producing

written descriptions that served as the primary data input for the AI. In the second cycle, these descriptions were submitted to ChatGPT using structured prompts requesting narrative sequence recommendations; the prompt design was refined across three separate sessions to improve output relevance. In the third cycle, the AI-generated sequence was evaluated against the director's aesthetic judgment, with each accepted or rejected recommendation recorded in a production log. In the fourth cycle, the edited film was reviewed as a completed artifact and the directorial decisions were systematically reflected upon. This reflective documentation forms the evidentiary basis for the findings reported in this study.

Ultimately, PLR in this study serves not only as a methodology but as a vehicle for navigating the evolving ecosystem of documentary creation within the digital landscape. It positions the filmmaker simultaneously as a scientist and a practitioner, seeking answers not just through conceptual frameworks, but through continuous, field-tested actions.

Phase 1: Data Collection

This phase involved 28 students enrolled in the fifth-semester Directing Practice course (Praktik Penyutradaraan) in the Film and Television Production program. Participants ranged from those with minimal prior fieldwork experience to those who had completed one previous production assignment. No standardized equipment was required: students used devices available to them personally, including smartphones (the majority), mirrorless cameras, and DSLR cameras(Bordwell, D., & Thompson, 2019). Recording duration was left open-ended, resulting in clips ranging from approximately 30 seconds to 4 minutes. Each student was asked to record at least one clip capturing an aspect of everyday life in Yogyakarta, with no predetermined subject, location, or aesthetic constraint. The resulting files were subsequently uploaded to a designated Google Drive repository.

Phase 2: Mapping and Analysis

The collected visual archives were processed through an analytical stage and categorized based on their respective titles. The following are the analytical descriptions of the 28 pieces of footage obtained:

1. Wide Shot. The courtyard of the Kauman Great Mosque in Yogyakarta, featuring visitors strolling and children cycling.
2. Medium Shot. The pedestrian area of Malioboro Street, Yogyakarta, focusing on three children sitting and eating ice cream amidst the crowd.
3. Wide Shot. A street vendor performing an animated dance while grilling satay, utilizing a "fire-trick" to create bursts of flame.
4. Wide Shot. A father and child playing in a park; the child stands on a brick raised bed and whispers to the father. A simple food and beverage stall is visible in the background.
5. Wide Shot. The Malioboro pedestrian area in front of storefronts, focusing on two men sitting on the floor playing chess. One man leans against a structural pillar, while the other sits on a *dingklik* (small wooden stool).

6. Medium Long Shot. A male snack vendor waiting for customers by the roadside, presenting his wares to the camera against a backdrop of passing vehicles.
7. Detail Shot. A gibbous moon in the night sky.
8. Full Shot. Two young individuals posing for the camera at Alun-alun Kidul at night, with toy vendors visible in the background.
9. Medium Shot. A fishmonger at a traditional market serving a customer.
10. Full Shot. Several women taking selfies with a water fountain in the foreground.
11. Wide Shot. A symmetrical roadside shot of an alley gate; the footage begins with a cyclist passing by, followed by a woman wearing a *kebaya* and *jarik* (traditional Javanese attire) exiting the gate.
12. Medium Shot. A female coffee vendor preparing a drink with a green thermos in the foreground while conversing with a customer.
13. Wide Shot. A riverside tourist spot where visitors are enjoying the view of the river and hills. The scene captures a family picnicking on a mat and young people taking photographs by the riverbank.
14. Close-up. A rotating disco ball reflecting light onto the ceiling.
15. Close-up (Self-video). A female student from the Film and Television Production program explaining her presence at a morning market to collect footage for a course assignment.
16. Full Shot. A man casting a net to catch fish.
17. Wide Shot. The summit of Mount Merapi.
18. Full Shot. Two monkeys with their offspring.
19. Wide Shot. A woman walking across the Alun-alun Kidul field using crutches, with the twin banyan trees in the background.
20. Wide Shot. Beachgoers walking on white sand, leaving visible footprints behind.
21. Full Shot. A group of people engaged in exercise, including running and walking.
22. Wide Shot. Visitors descending the stairs at the Imogiri Royal Cemetery, set against the backdrop of the steep, towering staircase.
23. Medium Full Shot. A male street performer playing guitar with a microphone and speaker at a traffic intersection, with vehicles passing in the background.
24. Extreme Wide Shot. A *Kuda Lumping* (traditional horse dance) performance at night, crowded with spectators. The angle is captured from behind the seated audience.
25. Wide Shot. A banyan tree at night with passing traffic in the background.
26. Medium Shot. A man recording himself via the reflection of a convex road mirror used as interior decoration, featuring a zoom-in.
27. Full Shot. A person riding a tall bicycle (*sepeda tinggi*) on city streets at night.
28. Timelapse. A sunset sequence.
29. Extreme Wide Shot (Bird's-eye view). Sand miners working to load sand onto a truck.

Phase 3: Directing (AI Collaboration via ChatGPT)

The practice-led research process involved constructing the narrative arc for the film titled "*Yogyakarta Hari Ini*" .The AI collaboration proceeded across three prompt iterations. In the first

iteration, screenshots of the full footage list—intentionally arranged in randomized, non-linear order—were submitted to ChatGPT with the prompt: "Based on these footage descriptions, suggest a narrative sequence for a short documentary about daily life in Yogyakarta." In the second iteration, the researcher refined the prompt by specifying desired emotional arc ("from intimate private moments to expansive public spaces, ending at dusk") and asking the AI to provide a rationale for each clip placement. In the third iteration, the AI was asked to propose two alternative sequences and explain the thematic logic of each. The AI's recommendations were documented and evaluated against the director's aesthetic judgment using a simple accept/revise/reject log for each clip placement. No human editor comparison was conducted in this study, which represents a limitation to be addressed in future research.

Figure 1
ChatGPT Screenshot

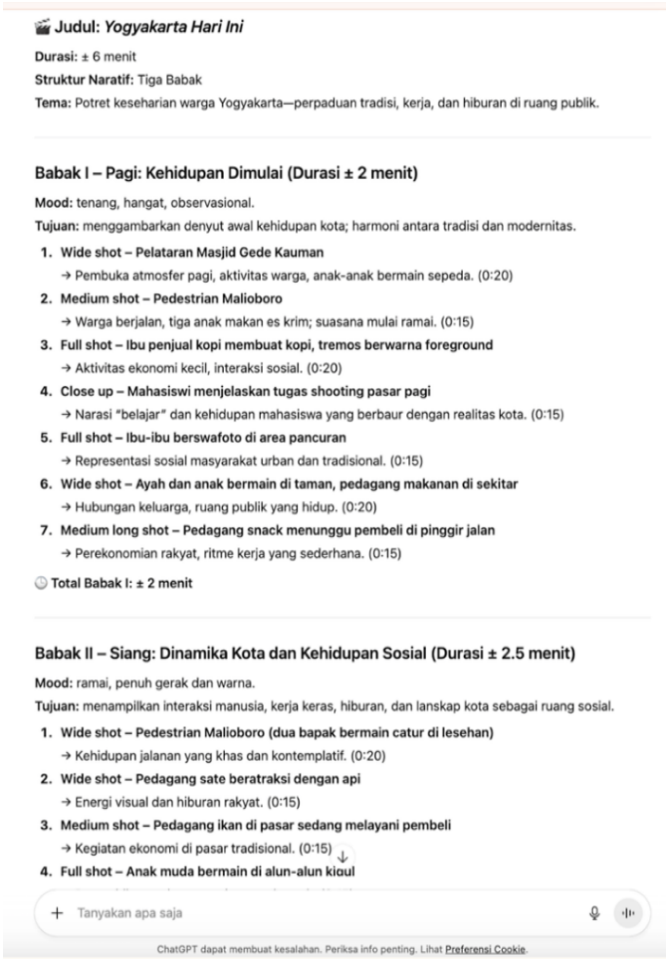
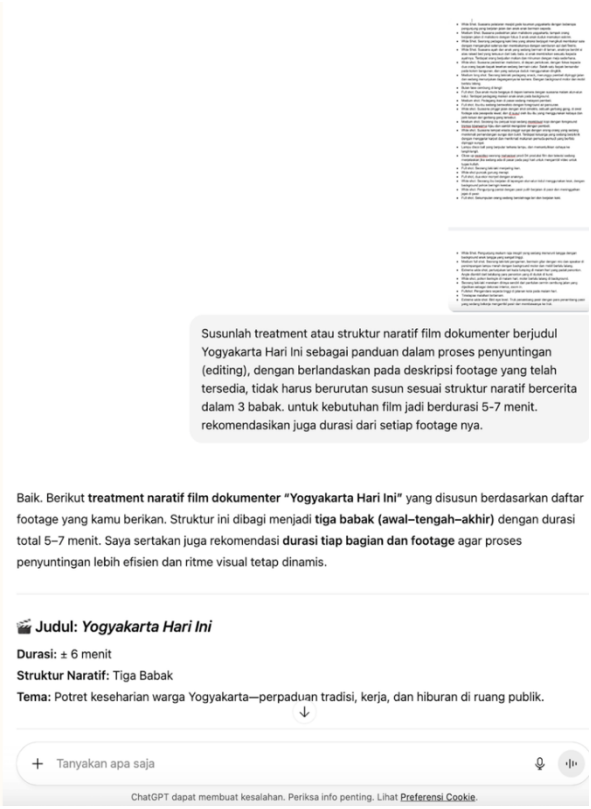


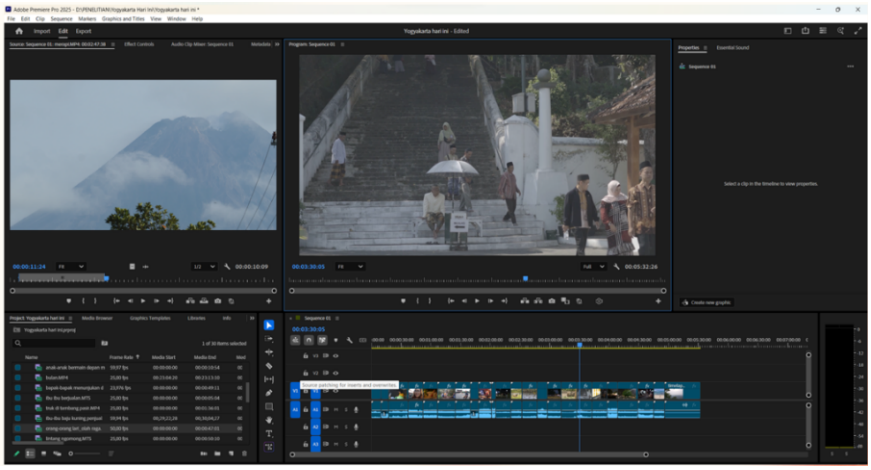
Figure 2
ChatGPT Screenshot



Phase 4: The Editing Process

This phase involved the assembly of the 28 pieces of footage based on the structural sequence recommended by the AI (ChatGPT).

Figure 3
Adobe Premiere Editing



Results

The research findings indicate that the collection of 28 footage clips recorded by students produces a highly diverse visual landscape of Yogyakarta, varying in recording styles, technical quality, and thematic focus. This diversity initially appeared as a fragmented reality lacking a central thread; a situation characteristic of the visual data deluge era. However, as the material was mapped, consistent thematic patterns emerged: the daily activities of citizens, dynamic public spaces, intimate interpersonal relationships, and a city rhythm fluctuating between tradition and modernity. These patterns formed the foundational framework for the documentary "*Yogyakarta Hari Ini*".

AI intervention during the directing and editing phases yielded several key findings. First, once the manual textual descriptions of all 28 clips were prepared, AI proved capable of categorizing and structuring those textual descriptions into a thematically coherent narrative sequence far more efficiently than traditional paper-edit methods. It is important to note that AI did not identify visual patterns autonomously from the footage itself; all visual analysis was performed manually by the researcher in Phase 2. AI's contribution was in the subsequent synthesis step: grouping the pre-described clips by thematic tendency (public space density, human movement, nature-to-city transitions) and proposing an initial sequence. While the AI-generated sequence was not entirely precise from an aesthetic standpoint, it provided a functional starting structure for narrative development. Within the context of PLR, the AI functioned as a "dramaturgical assistant," mapping potential narrative paths from the textual archive. Second, the process prompted the researcher to negotiate directorial intuition with the AI's data-driven arrangement, creating a productive dialogue between human creativity and computational pattern recognition. Third, the iterative nature of the prompt design process itself proved analytically significant: as the prompt was refined across three sessions—from a general narrative request, to one specifying an emotional arc, to one requesting multiple alternative sequences with thematic rationale—the quality and specificity of AI-generated recommendations improved substantially. This finding suggests that AI's utility in documentary narrative structuring is not fixed but highly responsive to the sophistication of human input, reinforcing its positioning as a collaborative tool rather than an autonomous agent.

During the editing stage, the AI-structured sequence accelerated the initial timeline assembly. However, the process also revealed substantive limitations. Of the 28 clip placements recommended in the primary sequence, approximately 9 (around one-third) required revision or rejection by the director. The most common failures involved: (1) transitions judged as too "mechanical," where the AI placed visually dissimilar shots in sequence without regard for tonal continuity; (2) disregard for atmospheric pacing, particularly the placement of high-energy performance footage (the Kuda Lumping clip) adjacent to intimate, static scenes; and (3) insensitivity to the symbolic weight of certain footage—for instance, the AI placed the crutches clip near the film's opening, where the director determined it held more resonance as a closing image. These corrections reveal that AI is effective as a structural scaffolding tool but consistently fails to account for the embodied, culturally-informed aesthetic judgment that

documentary directing requires. The creative process relies not only on sequential logic, but on intuitions of rhythm, visual texture, and meaning that remain beyond current AI capabilities. A fourth category of limitation also emerged during editing: the AI's tendency toward thematic closure that did not align with the director's intention to preserve a sense of open-endedness—a quality considered appropriate for a documentary representing the living, unresolved fabric of daily life in Yogyakarta. The AI consistently gravitated toward sequences that escalated in energy toward a clear climax and resolution, reflecting patterns prevalent in Western narrative traditions embedded within its training data. This finding has direct relevance to the cultural bias concerns (Bender, E. M.& Gebru, T.& McMillan-Major, A.& Shmitchell, 2021): the model's structural recommendations reflected a narrative logic more consonant with imported dramatic conventions than with the contemplative, accumulative structures characteristic of much Southeast Asian documentary practice. The director's interventions were therefore not merely aesthetic corrections but acts of cultural recontextualization—restoring a Javanese sensibility of rhythm and coexistence that the AI's recommendations had systematically displaced.

Overall, this study demonstrates that integrating AI into documentary directing, specifically when dealing with heterogeneous archives, can produce aesthetically accountable narrative arcs. PLR allows researchers to understand how technology transforms directorial workflows, viewing it not as a threat, but as a generative tool that expands the possibilities of contemporary documentary forms. Consequently, "*Yogyakarta Hari Ini*" serves not merely as a film, but as a prototype; a tangible example of how the documentary creation ecosystem in the digital age can evolve toward human-machine collaboration without compromising creative identity or sensitivity.

Discussion

This study demonstrates that the application of artificial intelligence in documentary directing based on heterogeneous visual archives is not only feasible, but also introduces new workflows relevant to today's digital landscape. Through a Practice-Led Research (PLR) approach, the filmmaking process for *"Yogyakarta Hari Ini"* confirms that AI can assist in organizing and proposing an initial structure for non-sequential footage, though this assistance is fundamentally dependent on prior human analysis. These findings are broadly consistent with Aufderheide's observation that AI tools in documentary post-production function most effectively as organizational accelerators rather than autonomous creative agents (Aufderheide, 2022). Similarly, the finding that approximately one-third of AI recommendations required directorial correction aligns with Amabile and Pratt's argument that computational creativity support tools augment rather than replace human creative judgment, particularly in domains where cultural context and emotional resonance are central criteria.(Amabile, T. M. & Pratt, 2016)

Nevertheless, the most critical finding lies in the inherent limitations of AI in this context. Machines lack aesthetic intuition, emotional depth, and sensitivity toward socio-cultural contexts that are fundamental to documentary filmmaking. This study further reveals a

dependency that prior literature does not always make explicit: AI narrative structuring tools cannot bypass manual human analysis; they operate downstream of it. The researcher must first interpret, describe, and contextualize every clip before AI can contribute meaningfully. This challenges any framing of AI as a labor-saving device in the visual analysis stage—it is more accurately a synthesis accelerator. This finding has important implications for how AI tools should be positioned in documentary pedagogy and professional practice: not as replacements for observational skill, but as tools that reward rigorous pre-analysis. Future research should address the absence of a comparative condition in this study—specifically, whether human editors working from the same textual descriptions produce sequences that are more, less, or equivalently coherent to the AI-generated ones, and across what criteria. A longitudinal study tracking how student filmmakers develop their use of AI across multiple productions would also yield important data on whether reliance on AI structuring accelerates or inhibits the development of independent dramaturgical judgment.

A third layer of significance concerns the pedagogical implications of AI-assisted directing within a formal film education context. The present study was embedded in a fifth-semester Directing Practice course at Institut Seni Indonesia Yogyakarta, positioning PLR not merely as a research methodology but as a curricular framework. The decision to integrate ChatGPT into an assessed creative workflow raises questions about how film schools should structure AI literacy alongside traditional craft competencies. As AI tools become increasingly embedded in professional documentary workflows globally—from AI-enabled transcription guiding assembly edits of lengthy interview material, to AI-generated assembly cuts replacing assistant editors in reality television production(International Documentary Association, 2024)—institutions training the next generation of documentary practitioners face an urgent curricular challenge. The prototype developed in this study suggests one viable pedagogical model: AI is introduced as a tool for narrative synthesis only after students have completed rigorous manual visual analysis, ensuring that the foundational observational and interpretive competencies of documentary practice are not bypassed. This sequencing is critical: it positions AI as a reward for thorough human labor, rather than a shortcut that circumvents it. Film programs across Indonesia and the broader Southeast Asian region may find this model replicable and adaptable to their own production contexts.

Finally, this study contributes to an emerging discourse on what might be termed post-auteur documentary practice in the digital landscape. The concept of the documentary director as sole creative authority has been progressively eroded by participatory, collaborative, and now computational filmmaking paradigms. Yet the findings of this study suggest that the director's role has not diminished so much as it has transformed: from primary recorder to curatorial architect, and now to critical negotiator between human sensibility and machine logic. The approximately two-thirds of AI sequence recommendations accepted by the director in this study were accepted not passively but through active evaluation against a set of embodied aesthetic criteria that the AI could not itself articulate or replicate. This negotiation may constitute a new mode of documentary authorship specific to the digital landscape era. It is a

mode in which the director's expertise is exercised not in the moment of recording, nor entirely in the moment of editing, but in the iterative and reflective space between human intention and algorithmic proposal. As the broader Indonesian film industry continues its rapid engagement with AI tools across production, visual effects, and distribution (Yulisman, 2025), understanding this negotiation space will be essential for practitioners, educators, and scholars alike who seek to ensure that technological adoption enhances rather than homogenizes the cultural specificity of Indonesian documentary cinema.

Conclusion

This research affirms that AI is not a threat to documentary directing; instead, it serves as a creative tool capable of expanding the possibilities of cinematic forms, methods, and aesthetics. The resulting methodological prototype is valuable not only within academic contexts but is also applicable to contemporary documentary productions facing a deluge of visual archives. This human-machine integration represents a logical step in the evolution of Indonesian documentary practice, while simultaneously marking a new phase in the way we interpret, assemble, and represent contemporary reality. The four iterative cycles of Practice-Led Research conducted in this study—from manual clip analysis through prompt refinement, directorial evaluation, and reflective documentation—together constitute a replicable workflow model that can be adapted by other institutions and practitioners working with heterogeneous visual archives. The model is particularly pertinent for participatory documentary projects, citizen journalism compilations, and multi-contributor archive-based productions, where the volume and diversity of material routinely exceed the capacity of conventional paper-edit processes.

Three specific directions for future research follow from this prototype. First, the study should be replicated with a comparative human-editor condition: a second group of editors should be given the same 28 textual clip descriptions and asked to construct narrative sequences without AI assistance, allowing for direct comparison of sequencing logic, thematic coherence, and cultural sensitivity. Second, the methodological framework should be extended to documentary projects involving non-urban and non-Javanese contexts within Indonesia, to test whether the AI's structural limitations in cultural attunement remain consistent across different regional aesthetics and narrative traditions. Third, the prototype warrants pilot implementation within formal film education curricula at multiple Indonesian institutions, with assessment instruments designed to measure whether the PLR-AI integration model accelerates or inhibits the development of students' independent directing judgment over time. Together, these extensions would strengthen the evidential basis for the human-machine collaborative model proposed here, and would contribute to a distinctly Indonesian body of knowledge on AI-assisted documentary practice—one that speaks from and to the cultural specificity of this archipelago's rich and heterogeneous visual traditions.

Author Contribution Statement

Diyah Verakandhi: Conceptualization and Research Design; Data Curation and Investigation; Methodology; Project Administration; Writing – Original Draft; Writing – Review & Editing.
Zahrina Zatadini: Methodology; Writing – Review & Editing; Validation.
Gregorius Arya Dhipayana: Formal Analysis and Visualization; Writing – Review & Editing.

References

Amabile, T. M. & Pratt, M. G. (2016). The dynamic componential model of creativity and innovation in organizations: Making progress, making meaning. *Research in Organizational Behavior*, *36*, 157–183. https://doi.org/https://doi.org/10.1016/j.riob.2016.10.001

Aufderheide, P. (2022). *Documentary film: A very short introduction (2nd ed.)*. Oxford University Press.

Bender, E. M.& Gebru, T.& McMillan-Major, A.& Shmitchell, S. (2021). On the dangers of stochastic parrots: Can language models be too big? In *2021 ACM Conference on Fairness, Accountability, and Transparency* (pp. 610–623). https://doi.org/https://doi.org/10.1145/3442188.3445922

Bordwell, D., & Thompson, K. (2019). *Film Art: An Introduction (12th ed.)*. McGraw-Hill.

Dean, R. T. & Smith, H. (2009). *Practice-led research, research-led practice in the creative arts*. Edinburgh University Press.

Doc Film Academy. (2026). AI in documentary filmmaking: Tools, ethics & tips. Retrieved from https://www.docfilmacademy.com/blog/ai-filmmaking

International Documentary Association. (2024). The synthesis: AI is everywhere, even in your documentary. Retrieved from https://www.documentary.org/column/synthesis-ai-everywhere-even-your-documentary

International Documentary Association (IDA). (2024). The synthesis: AI is everywhere, even in your documentary.

Manovich, L. (2013). *Software Takes Command*. Bloomsbury Academic.

Nichols, B. (2017). *Introduction to documentary (3rd ed.)*. Indiana University Press.

Renov, M. (2004). *The Subject of Documentary*. University of Minnesota Press.

Steyerl, H. (2015). *Duty Free Art: Art in the Age of Planetary Civil War*. Verso.

Strubell, E. & Ganesh, A. & McCallum, A. (2019). Energy and policy considerations for deep learning in NLP. In *The 57th Annual Meeting of the Association for Computational Linguistics*. https://doi.org/https://doi.org/10.18653/v1/P19-1355

Verakandhi, D. (2024). *Dokumenter Persalinan sebagai Representasi dari Realitas*. Institut Seni Indonesia Yogyakarta.

Yulisman, L. (2025). Indonesia's film industry embraces AI to make Hollywood-style movies for cheap. Retrieved from https://restofworld.org/2025/indonesia-ai-movies/