Proceedings of the 3rd International Conference on Machine Learning and Automation
DOI: 10.54254/2755-2721/2025.28119

# *Algorithmic Storytelling and Cinematic Narrative: A Comparative Study of AI-Generated Screenplays and Contemporary Auteur Cinema*

**Danqing Liu[1], Yuanchen Li[2*]**

*[1]The Hang Seng University of Hong Kong, Hong Kong, China*
*[2]Sapienza University of Roma, Rome, Italy*
**Corresponding Author. Email: rara481846778@gmail.com*

***Abstract.*** With the advancement of generative artificial intelligence, AI-based text generation has been increasingly applied to the domain of screenplay writing, raising critical questions about whether algorithmic storytelling can embody literariness, cultural expression, and philosophical depth. This study uses Life of Pi as a case and constructs two AI-generated screenplay samples (theme-driven and adaptation-driven) to compare systematically with Ang Lee's directorial version. Methods include narrative structure modeling, thematic weight analysis, symbolic language density computation, and philosophical abstraction measurement. A multidimensional comparison across narrative coherence, thematic focus, linguistic tension, and cultural depth is conducted, complemented by blind expert interviews involving five specialists to evaluate literary expressiveness from a humanistic perspective. The results show that while AI scripts perform well in structural control and thematic identification, they lag behind auteur-driven screenplays in philosophical abstraction, symbolic system construction, and aesthetic articulation. The study concludes that current AI systems are not yet capable of independently producing screenplays with humanistic depth but can function as effective tools in generating genre-oriented drafts.

***Keywords:*** AI screenplay generation, Auteur cinema, Literariness, Textual comparison, Narrative construction

## 1. Introduction

As generative AI technologies are progressing at a fast-paced rate, algorithmic text generation has increasingly been incorporated into the screenwriting practice. Recent models such as GPT-4 and ScriptWriter API are shown to create coherent narratives and genre-centered plot lines, presenting screenwriters with new means of creating drafts and conceptualizing story structure [1]. There has, though, been a persistent challenge, AI-generated screenplays tend to lag behind in expressing literary quality, intellectual depth, and cultural subtlety—essential dimensions traditionally maintained by author-centric cinema. In spite of past research having focused predominantly on the logic of structure and generative parsimony, scant consideration has accrued to the degree to which

© 2026 The Authors. This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (https://creativecommons.org/licenses/by/4.0/).

DOI: 10.54254/2755-2721/2025.28119

AI may mimic aesthetic and symbolical value embedded in human-created film narratives [2]. In this study, *Life of Pi* serves as a case to explore these tensions, creating two AI-driven screenwriting versions (thematically-oriented and adaptation-oriented) and comparing them systematically with Ang Lee's directing vision via semantic modeling, narrative cartography, and symbolical language detection. The ambition involves understanding the limitations and potential of literary expressiveness embedded in the prevailing frameworks of algorithmic storytelling.

## 2. Literature review

### 2.1. Evolution of AI narrative systems

More recent innovations in AI narrative systems employing large language models have greatly improved machine competency in creating plots, building characters, and linguistic flow. Programs such as GPT-4 now effectively mimic the three-act structure and generate dialogue sequences consonant with genre conventions [3]. Most of these systems, however, are fine-tuned to optimize structural competency and probabilistic fluency more than thematic resonance or symbolic articulation. Most of the literature emphasizes enhancing narrative coherence and genre conformity but neglects whether AI may create narratives projecting complex human experience or cultural specificity [4]. These systems then end up recreating familiar narrative shapes instead of creating original aesthetic configurations and thus suffer semantic flattening when confronted with abstract, philosophical, or metaphorical materials. As such, literary assessment of AI storytelling demands analytical instruments more sophisticated even than structural verification, and these are particularly needed in measuring symbolic compactness and hermeneutic penetration.

### 2.2. Auteur cinema: signature, philosophy, and cultural mediation

In accordance with the auteur theory, the director becomes the main person behind the narrative style of the film. It is directors such as Ang Lee who become the quintessence of this theory because they put in their personal philosophy and trans-cultural identity into the work [5]. In *Life of Pi*, Ang Lee combines the concept of the existential questioning along with the mythic narrative visualizing in order to create a story narrative beyond just a functional story structure. The movies produced by auteurs are not expected to fit in any particular genre owing to their non-linear narrative, symbolic indeterminacy, and open readability that AI cannot provide [6]. It has been seen from comparisons made between the two that while AI creates a logic coherence and pattern structure, it cannot create the symbolic density and cultural specificity found in auteur films [6]. What AI lacks is the "creative deviation" aspect that mimics real life and emotions.

### 2.3. Literariness in screen narrative

The literariness of screenwriting makes aesthetic beauty applicable to abstraction, symbolism, and multiple interpretations. Based on the poetic function from the Jakobson's perspective, writerly texts from the Barthes' point of view, as well as semiotic narrative approach, literariness as a register involves linguistic, cultural, and interpretive factors [7]. In movies, this phenomenon takes the form of metaphors, symbolic interlayers, structural ambiguity, and self-reflexivity. The modern AI algorithms, focused almost exclusively on statistical pattern recognition, create only superficial coherence and do not have any capacity to create flexible and symbolically rich texts. Being incapable to perceive cultural abstraction and complexity, AI-generated stories are always based on

DOI: 10.54254/2755-2721/2025.28119

direct interpretation and predictable meaning [8]. Thus, closing the possibility for multiple readings, AI narratives can be evaluated in terms of syntactic or structural coherence only. However, for assessing literary quality of AI-generated narratives, the semiotic and cultural-pragmatic analysis to determine whether AI texts have multivalence and resonance of human-generated literary texts is required.

## 3. Methodology

### 3.1. Sample and script generation

In the context of this study, *Life of Pi*, a film by Ang Lee, is chosen as the focal case of the comparative screenplay generation. The design of a two-track AI generation process is performed. First, a thematic screenplay is developed on the basis of GPT-4, employing semantic triggers such as "faith," "survival," "drift," and "illusion." The model is guided with structural pointers and emotion labels to develop a three-act screenplay in about 6000 words [9]. Second, an adaptation screenplay is generated based on the segmentation of the original literary work and its input into GPT-4, supported by LangChain-summarized narrative segments and the customized event extraction module. The training process of the latter is based on screenplay datasets from IMDb and the movies that won Academy Awards in order to guarantee the style consistency of the generated screenplay with the cinematic language. The output comprises three screenplay variants – AI original, AI adaptation, and the original director's version.

### 3.2. Analytical dimensions and evaluation metrics

To evaluate the narrative and literary quality of AI-generated versus director-authored screenplays, five quantifiable dimensions are modeled using computational linguistic techniques [10].

First, Narrative Coherence (NC) is measured by event alignment ratio:

$$NC=\frac{N_{aligned}}{N_{total}} \tag{1}$$

Where $N_{aligned}$ refers to the number of events in the text that successfully match standard narrative structure elements (such as conflict, turning point, and climax), and $N_{total}$ represents the total number of narrative events in the text.

Second, Character Depth (CD) evaluates psychological complexity via internal conflict density:

$$CD=\frac{\sum_{i=1}^{n}C_i}{n} \tag{2}$$

Where n is the number of main characters in the text and $C_i$ denotes the frequency of internal conflict, emotional fluctuation, or self-contradiction presented by character i in the screenplay.

Third, Thematic Density (TD) employs a weighted TF-IDF formula to assess thematic focus:

$$TD_k=\frac{tf_k\cdot\log\left(\frac{N}{df_k+1}\right)}{L} \tag{3}$$

Where $tf_k$ indicates the term frequency of the kth thematic keyword such as faith or survival in the text, N is the total number of documents in the corpus, $df_k$ is the number of documents containing that keyword, and L is the total number of words in the text.

DOI: 10.54254/2755-2721/2025.28119

Fourth, Symbolic Language Density (SLD) is calculated as symbolic phrase frequency per sentence:

$$SLD=\frac{N_{symbolic}}{N_{sentences}} \tag{4}$$

Where $N_{symbolic}$ refers to the number of symbolic expressions identified through semantic annotation such as ocean symbolizing faith or tiger symbolizing inner fear and $N_{sentences}$ is the total number of sentences in the text.

Fifth, Philosophical Abstraction Tension (PAT) utilizes cosine distance in BERT embeddings:

$$PAT=\frac{1}{|P|}\sum_{p\in P}\text{cosine_distance}\left(v_p,v_c\right) \tag{5}$$

Where P is the set of philosophical keywords such as truth existence and faith $v_p$ is the embedding vector of the keyword in the BERT semantic space and $v_c$ is the average contextual vector in which the word appears.

These formulas enable the structured comparison of symbolic density, semantic abstraction, and narrative architecture across the three screenplay types.

### 3.3. Expert interviews and qualitative assessment

In this research, five expert reviewers including screenplay writers, literary critics, and narrative experts were involved in conducting the expert interviews. The approach used for this study includes semi-structured interview and blind review carried out using online video conferencing technology. The interviews were 60 minutes each. The objective is to carry out qualitative validation from textual meaning, symbol construction, and cultural depth perspectives. Each reviewer was allotted randomly one AI original screenplay, one AI adapted screenplay, and one director screenplay. The interview content is shown in Table 1.

Table 1. Interview question table

| Analytical Dimension | Guiding Question |
|---|---|
| Authorial Perspective | Can you perceive any distinct creative standpoint or value expression in the text? |
| Symbolism and Metaphor | Do you observe any symbolic or non-literal expressions in the text? Can you give examples? |
| Cultural or Philosophical Position | Does the text reflect any cultural identity, religious reflection, or philosophical theme? |
| Emotional and Aesthetic Expression | Does the text evoke any emotional reaction or contain a distinct linguistic aesthetic? |

Expert opinions were transcribed and thematically coded with consent. Through semantic clustering by NVivo 14, this research gathered similar and differing perspectives in order to formulate an evaluation framework for AI screenplay based on humanistic perspectives. Findings indicate that whereas AI screenplays are coherent in terms of their narrative consistency, there is a lack of symbolism and culture as well as emotional aesthetics within the AI screenplay [11]. There is

DOI: 10.54254/2755-2721/2025.28119

an indication of the usage of functional language in place of authorial language, whereby symbolic content seems to follow templates.

## 4. Results

### 4.1. Narrative structure and thematic consistency

In terms of the structural characteristics of AI-generated screenplays in relation to narrative coherence, the results obtained from the quantitative study of the three kinds of screenplays reveal their distinctive traits. According to the results depicted in Figure 1, the value of narrative coherence index for the theme-driven AI screenplay equals 0.78, for the adaptation-driven AI screenplay equals 0.82, and for the original screenplay created by the director equals 0.91. In the case of the measurement of thematic density, the high level of frequency of keyword occurrence is demonstrated by AI-generated screenplays. Thus, the theme-driven script has the TF-IDF weighted values of 0.156 and 0.143 for the keywords of faith and survival, whereas the director's screenplay has the corresponding values of 0.089 and 0.092. Consequently, it can be claimed that AI is able to create coherent stories that have consistent narrative and the ability to capture key themes of the story. However, the highly concentrated nature of the themes in AI-generated texts reveals some mechanical features of semantic expression, which lack natural variability typical for human-written literature. In comparison with the theme-driven screenplay, the adaptation-driven script demonstrates higher integrity in its structure.

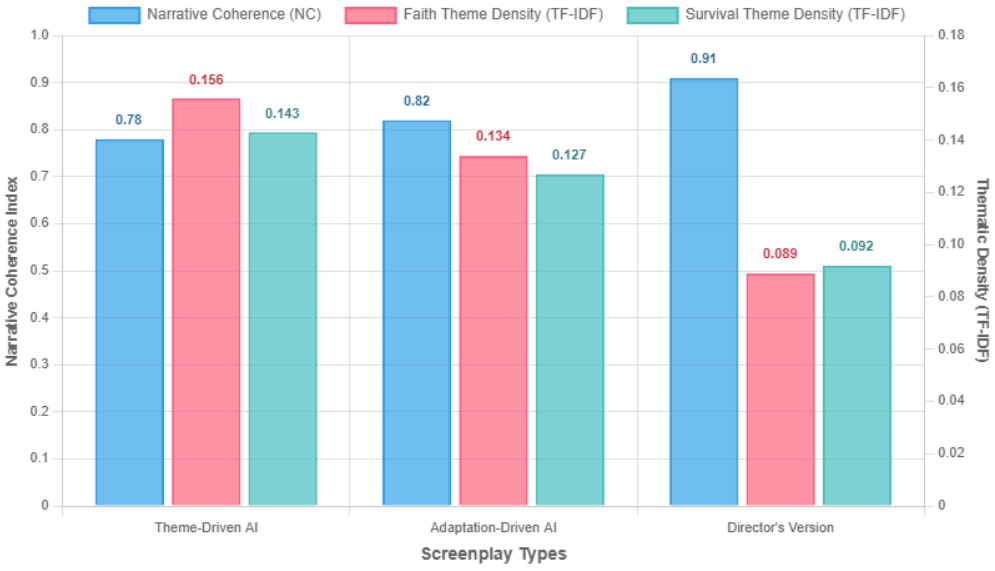



Figure 1. Comparative analysis of narrative coherence and thematic density

### 4.2. Philosophical and literary expression

In the measurement of symbolic language density and philosophical abstraction tension, the three types of screenplays present significant differences as shown in Table 2. The director's version reaches 0.31 symbolic expressions per sentence in symbolic language density, while the AI-generated scripts contain only 0.18 in the theme-driven version and 0.22 in the adaptation-driven version. More critically, in the BERT embedding distance analysis of philosophical abstraction

DOI: 10.54254/2755-2721/2025.28119

tension, AI screenplays show lower semantic complexity, with an average cosine distance of 0.67, which is substantially lower than the director's version at 0.89.

Table 2.

| Screenplay Type | Symbolic Language Analysis | | Philosophical Abstraction | | Expert Assessment | |
|---|---|---|---|---|---|---|
| | SLD (per sentence) | Metaphor Frequency | PAT (BERT distance) | Abstract Concepts | Cultural Depth (1-5 scale) | Aesthetic Quality (1-5 scale) |
| Theme-Driven AI | 0.18 | 23 | 0.67 | 42 | 2.1 | 2.3 |
| Adaptation-Driven AI | 0.22 | 31 | 0.71 | 38 | 2.4 | 2.6 |
| Director's Version (Ang Lee) | 0.31 | 67 | 0.89 | 78 | 4.2 | 4.5 |

Interview outcomes from experts give additional support to the findings obtained in the course of the quantitative study. According to four out of five experts, there are noticeable limitations in the way AI screenplays reflect culturally and emotionally meaningful aesthetic elements. Though the use of symbolism can be identified in screenplays, the symbolic elements are rather templated and do not possess the organically created context. As far as philosophy is concerned, AI tends to express it through explicit statements instead of using metaphors. Even though such an approach to the construction of texts makes the content easily understandable, the openness of the text and its aesthetic tension are diminished.

## 5. Discussion

Through comparative study of three kinds of screenplays in terms of structural logic, thematic construction, symbolic regimes and philosophy, this study finds that in terms of the technical capabilities of generating AI screenplay in the area of narrative conciseness and structure manipulation, but also critically deficient in literariness and cultural construction. The AI script has shown itself to be consistent and efficient in template-based identification of themes and narrative logic, which is well suited to imitating the genre-based pattern, but this is accompanied by semantic uniformity and symbolic compression, thus the incapacity to cope with ambiguity, cultural meaning, or emotional depths. Expert interviews further reinforce this point that although AI script may imitate the process of human script writing, they lack the active role and cohesion in the purpose of creation, linguistic tension, or worldviews.

## 6. Conclusion

Using *Life of Pi* as a case study, this paper constructed both theme-driven and adaptation-driven AI screenplays and compared them with the director's original version to systematically examine the boundaries and potential of generative AI in narrative construction and literary expression. The results show that the screenplays written by artificial intelligence are marked with a certain level of technological maturity in terms of structure, pacing, and theme, thus being appropriate for genre-based story telling at an initial level. Yet, regarding the symbolic systems building, cultural

DOI: 10.54254/2755-2721/2025.28119

grounding, and philosophical semantic representation, AI-written texts manifest a certain level of flatness and mechanization without any artistic and intellectual profundity. The qualitative interview confirms AI's lack of presence of the author, emotions, and language artistry. In the future studies, there should be more attention paid to multimodal input, value-oriented modeling, and cross-cultural learning of hybrid narrative models.

## Contribution

Danqing Liu and Yuanchen Li contributed equally to this paper.

## References

[1] Mirowski, P., et al. (2023). Co-writing screenplays and theatre scripts with language models: Evaluation by industry professionals. *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*.
[2] Lin, Z., et al. (2025). R$^2$: A LLM based novel-to-screenplay generation framework with causal plot graphs. *arXiv preprint arXiv:2503.15655*.
[3] Tang, Y., et al. (2025). Understanding screenwriters' practices, attitudes, and future expectations in human-AI co-creation. *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems*.
[4] Xu, Y. (2025). Balancing creativity and automation: The influence of AI on modern film production and dissemination. *arXiv preprint arXiv:2504.19275*.
[5] Elias, S., et al. (2025). Rethinking literary creativity in the digital age: A comparative study of human versus AI playwriting. *Humanities and Social Sciences Communications, 12*(1), 1–11.
[6] Serena, A. (2025). *The impact of generative AI on the film industry: Opportunities, challenges, and ethical implications* [Doctoral dissertation, Politecnico di Torino].
[7] Prabowo, B. A., & Asmarani, R. (2025). Generative literature: The role of artificial intelligence in the creative writing process. *Allure Journal, 5*(1), 1–9.
[8] Holopainen, J., et al. (2025). Infinity book: Speculating literary expressions in the age of generative AI. *Proceedings of the 2025 ACM Designing Interactive Systems Conference*.
[9] Dayo, F., Memon, A. A., & Dharejo, N. (2023). Scriptwriting in the age of AI: Revolutionizing storytelling with artificial intelligence. *Journal of Media & Communication, 4*(1), 24–38.
[10] Vainikka, E., Soronen, A., & Kallio, S.-M. (2025). Is prompt engineering the future of screenwriting? Views of professional screenwriters and commissioners about the impact of AI technologies on their profession. *Critical Studies in Television, 20*(1), 54–72.
[11] Anantrasirichai, N., & Bull, D. (2022). Artificial intelligence in the creative industries: A review. *Artificial Intelligence Review, 55*(1), 589–656.