



AI AS A STORYTELLER: A Comparative Narrative Analysis of ChatGPT and Claude

ALMA LLERENA-FERNÁNDEZ ¹

alma.llerena@urjc.es

¹Universidad Rey Juan Carlos, España

KEYWORDS	ABSTRACT
Artificial intelligence AI-generated narratives Screenplay analysis Narrative structure Algorithmic bias Computational narratology Generative language models	<i>This study examines AI-generated narratives through the analysis of film scripts, focusing on how different large language models construct narrative structure, characters, and the presence of bias under controlled conditions. Using a mixed-method experimental design, eight scripts generated by ChatGPT 5.1 and Claude Sonnet 4.5 were analyzed across two narrative genres (drama and thriller) and under two generation conditions: demographic neutrality and identity permissiveness. The results indicate that identity-related conditions influence the narrative configuration of the scripts and reveal distinct narrative signatures across models. While ChatGPT tends toward more conventional structures and the reproduction of implicit narrative patterns, Claude increases structural, psychological, and relational complexity, albeit with a higher incidence of formal inconsistencies.</i>

Recibido: 25/ 03 / 2026

Aceptado: 18/ 06 / 2026

1. Introduction

Over recent decades, the screenplay has been the subject of extensive theoretical development that consistently recognizes its central function as the narrative architecture of the audiovisual work (Aumont et al., 2019; Bang, 2022). Beyond its consideration as a technical document, the screenplay structurally organizes the arrangement of narrative events, the construction of characters, the progression of conflict, and the articulation of time and space, thus functioning as a preliminary device that decisively conditions the cinematic experience (Field, 2005; McKee, 2009; Sánchez Escalonilla, 2008; Seger, 2017). This conception has consolidated the screenplay as an autonomous analytical object, amenable to systematic and comparative narrative analysis (Bang, 2022; Barthes & Duisit, 1975;).

From this perspective, studies on audiovisual narration have highlighted the persistence of relatively stable structures in fiction cinema, particularly in feature-length films, where act-based organization, dramatic causality, and the centrality of conflict continue to operate as dominant structural principles (Bordwell et al., 2020; Cuadrado Alvarado, 2017; Field, 2005; García Serrano, 2019; McKee, 2009). However, these models do not function as closed formulas, but rather as flexible frameworks that allow for variation, combination, and deviation depending on genre conventions, production contexts, and authorial decisions (Dancyger & Rush, 2002; Vanoye, 1996; Yorke, 2014). Precisely for this reason, narrative structure has been consolidated as a key analytical tool for the comparative study of screenplays, as it enables the identification of regularities, limits, and strategies of narrative organization.

Complementarily, screenplay theory has emphasized the central role of characters as drivers of dramatic action, understanding that conflict and its development are articulated through the desires, decisions, and transformations of narrative agents (McKee, 2022; Seger, 2017; Zimmerman, 2018). Actantial configuration, character arcs, and relationships between characters acquire particular relevance in genres such as drama and thriller, where the tension between internal and external conflict structures narrative progression (Altman, 2010; Chatman, 1980). Added to this is the management of narrative space and time, through scene segmentation, the use of ellipsis, and rhythmic organization, as fundamental elements for the coherence and legibility of the narrative universe (Carmona, 2016; Casetti & Di Chio, 2017; Martin, 2008).

The emergence of generative language models based on artificial intelligence introduces a scenario that directly challenges some of the traditional assumptions of screenplay analysis. The capacity of these systems to produce extended, coherent, and formally structured texts has reactivated debates surrounding authorship, creativity, and the very nature of narrative construction, compelling a reconsideration of the extent to which screenplay conventions can be reproduced, simulated, or tensioned by non-human generative processes (Manovich, 2020; Rajas & Gértrudix, 2025).

In recent years, various studies have addressed AI-assisted screenwriting from perspectives focused on creative collaboration, the exploration of narrative variations, or the optimization of writing processes. From a cultural and aesthetic approach, Manovich (2020) situates generative production within contemporary logics of computational creativity, emphasizing that these systems do not generate works *ex nihilo* but operate through the recombination of previously learned patterns. Within the specific domain of screenwriting, Viljoen (2024) and Cake (2025) analyze the use of generative models as assistants in early stages of narrative development, highlighting both their operational potential and persistent limitations in structural control and global coherence across generated texts.

Nevertheless, and despite the growth of this body of literature, most existing studies have focused on discussing the creative possibilities of artificial intelligence or on providing general evaluations of the quality of generated texts, without systematically examining how these systems manage the complex narrative structures characteristic of feature-length screenplays (Bender et al., 2021; Manovich, 2020). In particular, empirical research that comparatively analyzes the construction of narrative structure, character development, and spatiotemporal organization in screenplays generated by different language models under controlled narrative conditions remains limited, a shortcoming repeatedly identified in recent scholarship on AI-assisted audiovisual writing (Rajas & Gértrudix, 2025).

From an approach directly linked to audiovisual narration, Rajas and Gértrudix (2025) specifically point to the difficulty generative systems face in sustaining complex and stable narrative architectures

in extended texts, underscoring the need for studies that examine these screenplays not only as creative products but also as narrative structures susceptible to formal analysis. In this sense, the reproduction of recognizable screenplay conventions by artificial intelligence does not in itself guarantee narrative coherence, but rather constitutes a starting point from which to interrogate its structural limits.

In this context, the present study intervenes directly in this gap through a systematic and comparative narrative analysis of screenplays generated by artificial intelligence. The research focuses specifically on examining narrative structure, character construction, and the organization of narrative space and time, as well as on identifying errors and representational biases, in order to assess the extent to which generative models are capable of sustaining the formal and narrative demands of feature-length fiction. In contrast to approaches centered on subjective evaluations of creativity or originality, this study adopts a rigorous analytical framework grounded in consolidated proposals linked to screenplay practice and analysis (Bordwell et al., 2020; Field, 2005; McKee, 2009; Seger, 2017; Tobias, 1993; Vogler, 2020), thereby providing empirical evidence to a debate that has thus far been dominated by theoretical or exploratory approaches.

1.1. Objectives

The general objective of this study is to systematically analyze the capacity of different artificial intelligence models to generate cinematic screenplays, with particular attention to their adherence to the narrative conventions of fiction cinema and to the differences derived from the generative model, narrative genre, and generation conditions.

More specifically, the study pursues the following objectives:

- To examine how artificial intelligence models structure cinematic narratives in the drama and thriller genres, focusing on the organization of narrative structure, the type of conflict, narrative closure, and their alignment with consolidated narrative models.
- To analyze the construction of AI-generated characters, paying particular attention to actantial definition, character arcs of transformation, and relational dynamics, as well as their coherence with genre-specific conventions.
- To identify patterns in the use of narrative space and time, evaluating the degree of clarity, continuity, and rhythm control in the generated screenplays, as well as the deployment of specific temporal devices.
- To compare the narrative strategies adopted by different artificial intelligence models in terms of structural complexity, actantial density, and formal control of the screenplay, relating these aspects to professional formatting standards and the internal coherence of the narrative.
- To detect and describe the presence of narrative, formal, and coherence-related errors, as well as the emergence of representational biases, in order to assess the current limitations of AI-generated screenwriting.

Taken together, these objectives enable an analytical and comparative approach to AI-generated screenwriting, aimed at understanding not only the narrative outcomes produced, but also the underlying logics and constraints that shape the algorithmic generation of cinematic narratives.

2. Methodology

To achieve the stated objectives, an experimental methodology with a mixed-methods approach was designed, combining qualitative procedures of observation, identification, analysis, and comparison of concepts related to the construction of cinematic narratives with descriptive quantitative metrics. The study was structured around two key stages: the design of the prompts used to generate the screenplays that constituted the sample, and the development of a specific analytical model for the examination of the resulting corpus.

The research was conceived from a comparative and controlled logic, transferring tools traditionally used in screenplay analysis to the study of algorithmically generated texts. In this regard, the screenplays produced by artificial intelligence models were considered narrative texts methodologically comparable to screenplays written by human authors and, therefore, amenable to

analysis through concepts associated with creative practice and consolidated analytical models in audiovisual narrative studies.

The study corpus consisted of eight screenplays generated specifically for this research using two advanced large language models: ChatGPT 5.1 and Claude Sonnet 4.5. In order to ensure the comparability of results, an experimental matrix was designed that combined two narrative genres, drama and thriller, with two differentiated generation conditions. Condition A, defined as demographic neutrality, stipulated that the model should not explicitly assign attributes related to gender, ethnicity, nationality, or sexual orientation to characters, except in cases where such traits emerged organically and were narratively necessary. Condition B, referred to as identity permissiveness, allowed the model to freely assign such attributes without explicit restrictions. This combination resulted in four screenplays generated per model, yielding a corpus balanced in terms of narrative genre, experimental condition, and generative system.

Screenplay generation was carried out through a system of standardized prompts designed to precisely control narrative production conditions and ensure comparability across the generated texts. Given that artificial intelligence models do not generate long-form narrative texts in a continuous manner, each screenplay was structured through a sequence of four consecutive prompts, formulated according to a progressive and cumulative logic that enabled continuous narrative development while maintaining coherence and avoiding rewrites or reformulations.

In all cases, the first prompt explicitly established the narrative genre, professional screenplay formatting instructions, and a set of shared narrative guidelines, such as the presence of two main characters, a primary plot and at least one interconnected subplot, as well as the requirement not to conclude the story in the initial stages of generation. The intermediate prompts were exclusively oriented toward narrative continuation, incorporating strict instructions regarding internal coherence and narrative continuity, while avoiding any form of correction, summarization, or reinterpretation of previously generated content. Finally, the last prompt requested the complete closure of the narrative arc and, in a controlled manner, a brief reflection by the model on potential biases present in the generated story.

The prompts also incorporated explicit clauses intended to constrain model behavior, such as prohibiting the use of external tools, searches, or clarification questions, in order to analyze the resulting narrative output exclusively on the basis of the system's internal capabilities. With respect to the experimental conditions, differentiated instructions were introduced depending on the degree of control over characters' identity attributes, allowing observation of how narrative construction varies under a framework of demographic neutrality versus identity permissiveness.

This prompt design made it possible to generate a methodologically homogeneous corpus, maintaining constant production conditions and ensuring that observable variations across the screenplays were attributable to the narrative behavior of the models, the genre, and the experimental condition, rather than to differences in instruction formulation.

To preserve methodological validity, technical generation parameters were held constant within each artificial intelligence model. No manual creativity adjustments were applied, and extended reasoning modes were not activated. In cases where the temperature parameter was not user-configurable, the system's default value (≈ 0.7) was employed, allowing for a balance between narrative coherence and creative diversity. This methodological decision was intended to analyze the standard narrative behavior of each system as accessible to advanced users and to enhance replicability by avoiding highly customized configurations.

Once generated, the screenplays were consolidated into independent documents and left entirely unedited, with no stylistic, narrative, or formal corrections, in order to preserve the original output of the models in full. Each screenplay was generated in independent sessions or chats, thereby maximizing variability within the same experimental condition in accordance with the stochastic nature of large language models.

Based on this material, an ad hoc narrative analysis model was developed, specifically designed to examine AI-generated screenplays from a structural, functional, and representational perspective. The model was structured around ten variables: *general characteristics*, *structure*, *characters*, *space and time*, *drama genre elements*, *thriller genre elements*, *errors*, *biases*, *graphic codes*, and *sound codes*.

Within the general characteristics variable, nine categories were included, addressing aspects considered fundamental in the generation of screenplays using artificial intelligence models. These included: word count, number of scenes, the proportion of dialogue to action, the number of main, secondary, and background characters, the total number of characters, adherence to standardized screenplay formatting, the inclusion of a title, and the number of macro-locations and micro-locations used in the construction of narrative spaces.

Complementarily, narrative locations were cataloged using a normalized spatial taxonomy. Lexical variants referring to the same functional space (for example, car, SUV, or another character's vehicle) were grouped into a single location category, prioritizing narrative function over surface description and facilitating comparative analysis across models.

For the *structure* variable, reference was made to major theoretical frameworks associated with audiovisual screenwriting and narrative structure analysis, drawing on the contributions of Altman (2010), Cruz (2024), Field (2005), McKee (2009), Sánchez-Escalonilla (2008), Seger (2018), Tobias (1993), Vogler (2020) and Yorke (2014). These approaches allow for the identification of fundamental structural points within the narrative and for the analysis of internal coherence, as well as the relationship between the screenplay and different narrative models, such as the classical three-act design, the five-act theatrical structure, or the hero's journey. Based on these frameworks, a set of twenty categories was defined for narrative structure analysis.

Within the *characters* variable, nine analytical categories were defined based on the theoretical contributions Field (2005), García Barrientos (2007), McKee (2022), Sánchez-Escalonilla (2008) and Seger (2017), aimed at examining how characters are constructed within the narrative. These categories allow for the analysis of characters' narrative functions, the presence and development of transformation arcs, characterization mechanisms, typological variety, and relational dynamics, as well as the presence of characters who, while not directly appearing in the action, play a relevant role in plot articulation.

To analyze how artificial intelligence articulates narrative *space and time*, six analytical categories were established based on the contributions of Carmona (2016) and Casetti and Di Chio (2017). These categories make it possible to examine the use of temporal devices such as flashbacks or ellipses, coherence in spatial and temporal organization, and the management of narrative rhythm in relation to story progression.

Regarding variables linked to genre-specific elements of *drama* and *thriller*, analytical categories were established based on the contributions of de Altman (1999), Bordwell et ál. (2020), Chatman (1978), McKee (2009) and Schatz (1981). These categories allow for the examination of the extent to which AI-generated screenplays adhere to the narrative, thematic, and formal conventions associated with each genre, as well as the identification of potential hybridizations or deviations from their characteristic traits.

In addition, twelve categories were incorporated into the *errors* variable, aimed at identifying different types of inconsistencies and failures that may occur in the creation and writing of AI-generated screenplays. Complementarily, ten categories were defined within the biases variable, focused on analyzing the assignment of identity attributes, the reproduction of narrative hierarchies, and the presence of stereotypes in character and conflict construction.

Finally, drawing on the contributions of Carmona (2016) and Casetti and Di Chio (2017), the model incorporated three categories dedicated to the analysis of *graphic codes* and two categories focused on *sound codes*, in order to observe the use of these expressive resources in narrative construction.

The variables and categories defined in the model were operationalized through procedures of identification, counting, and categorical classification, according to the nature of each variable. These criteria were applied homogeneously across the entire corpus to ensure internal analytical coherence and comparability among the generated screenplays.

The application of the analytical model was carried out systematically for each screenplay in the corpus, enabling comparative readings across texts, genres, experimental conditions, and generative systems, without introducing interpretive considerations related to results at this stage, which are addressed in the corresponding section.

3. Results

Within the *general characteristics* variable, the results show significant differences in screenplay length and in the segmentation of narrative locations depending on the generative model, genre, and experimental condition. As shown in Table 1, the screenplays generated by Claude systematically include a higher number of scenes than those produced by ChatGPT, both in drama and thriller and under both conditions, resulting in greater narrative and spatial fragmentation.

Table 1. Results of the distribution of locations.

	CGP CND A Drama	Claude CND A Drama	CGP CND B Drama	Claude CND B Drama	CGP CND A Thriller	Claude CND A Thriller	CGP CND B Thriller	Claude CND B Thriller
Domestic space	-	13	15	36	5	13	-	11
Institutional space	-	32	2	14	-	-	-	3
Functional public space	15	1	7	1	5	25	-	11
Open urban space (street / road)	-	4	5	14	3	15	3	7
Vehicle / transport	-	1	4	-	4	3	7	-
Enclosed space of threat / control	-	-	-	-	9	1	34	23
Other / unspecified	16	25	5	27	14	29	3	27
Total scenes (N)	31	76	38	92	40	86	47	82

Fuente: Own elaboration, 2026.

The analysis of location distribution also reveals consistent genre-based patterns. In dramatic screenplays, action is more concentrated in domestic and institutional spaces, especially in texts generated by Claude, whereas thrillers more frequently incorporate locations associated with transit and action and, notably, enclosed contexts of threat or control, particularly under Condition B. Condition A, for its part, tends toward more contained spatial segmentation and a higher presence of generic or unspecified spaces, whereas Condition B expands the repertoire of locations and the total number of scenes, increasing the complexity of the narratives' spatial universe.

Within the *general characteristics* variable, the data on number of characters, spatial segmentation, and maintenance of professional formatting show consistent differences between the generative models (cf. Table 2). Screenplays produced by Claude systematically exhibit a higher total number of characters, with particularly pronounced proliferation of secondary characters and extras, as well as a significant increase in the number of macro-locations and micro-locations, resulting in greater complexity and narrative fragmentation. This pattern is especially evident in thrillers and under Condition B.

Table 2. Results for the *general characteristics* variable.

	CGP CND A Drama	Claude CND A Drama	CGP CND B Drama	Claude CND B Drama	CGP CND A Thriller	Claude CND A Thriller	CGP CND B Thriller	Claude CND B Thriller
No. of main characters	2	3	2	2	2	4	1	4
No. of secondary characters	1	4	2	5	2	5	4	7
No. of extras	1	17	2	28	-	>40	1	>50
Total no. of characters	4	24	6	35	4	49	6	51
Maintains professional format	YES	NO	YES	NO	YES	Sometimes	YES	Sometimes
Macro-locations	4	10	13	19	14	29	8	30
Micro-locations	8	28	12	31	12	46	17	27
Includes a film title	NO	NO	NO	NO	NO	YES	NO	YES

Fuente: Own elaboration, 2026.

By contrast, screenplays generated by ChatGPT display more contained structures, both in the number of characters and in spatial diversity, consistently maintaining professional screenplay formatting. In Claude-generated texts, however, adherence to formal conventions is more irregular, especially in thrillers, where increased narrative scale is accompanied by a loss of formal control.

Overall, these results suggest that the observed differences in spatial segmentation, number of characters, and format control reflect differentiated narrative strategies across generative models. Whereas Claude tends to expand the narrative universe through the multiplication of spaces and characters, favoring narratives of greater density and fragmentation, ChatGPT appears to prioritize a more contained and stable organization, with stronger adherence to the formal conventions of the screenplay. This divergence intensifies in the thriller genre and under Condition B, where the expansion of the spatial and actantial repertoire is associated with greater loss of structural control, pointing to a direct relationship between narrative complexity and formal stability in AI-generated screenplay.

Alongside these quantitative differences, qualitative divergences in writing style were also identified between the screenplays produced by both models. ChatGPT-generated texts tend toward a simple and functional syntax, with a predominance of short sentences, limited subordination, and a uniform external narrative voice, resulting in a repetitive style oriented toward expository clarity. By contrast, Claude-generated screenplays exhibit greater syntactic complexity and lexical richness, with greater rhythmic variation and a more marked narrative voice that, in some cases, approaches the characters' interiority. This stylistic density contributes to a more literary mode of writing, although it is also associated with lower formal stability in maintaining professional screenplay format.

As shown in Table 3, results for the *structure* variable indicate a high degree of adherence to classical narrative models across the corpus, with consistent patterns depending on the generative model and narrative genre. All analyzed screenplays systematically follow a three-act structure, evidencing robust assimilation of the archplot (McKee, 2009) or paradigm (Field, 2005) as a basic organizational scheme by both artificial intelligence models.

Table 3. Results for the *structure* variable.

	CGP CND A Drama	Claude CND A Drama	CGP CND B Drama	Claude CND B Drama	CGP CND A Thriller	Claude CND A Thriller	CGP CND B Thriller	Claude CND B Thriller
Maintains plausibility	YES	YES	YES	YES	Sometimes	Sometimes	Sometimes	Sometimes
Follows a 3-act structure	YES	YES	YES	YES	YES	YES	YES	YES
Follows a 5-act structure	NO	YES	YES	YES	NO	YES	NO	YES
Follows the hero's journey structure	NO	YES	YES	NO	NO	YES	NO	YES
Type of main conflict	Internal	Internal	Internal	Internal	External	External	External	External
Type of ending	Closed	Closed	Closed	Closed	Closed	Closed	Closed	Closed
Presence of subplots	NO	YES	YES	YES	NO	NO	NO	NO

Fuente: Own elaboration, 2026.

Beyond this shared pattern, relevant differences emerge in the structural complexity adopted by each system. Whereas ChatGPT-generated screenplays tend to adhere more strictly to a single structural model, Claude-produced texts show a greater propensity to overlay multiple narrative schemes, more frequently incorporating elements of the five-act structure and the hero's journey, especially in dramatic screenplays and under Condition B. This combination suggests a narrative strategy aimed at increasing structural density.

With respect to genre, results confirm a clear correspondence between structure and type of conflict. In dramatic screenplays, internal conflict predominates consistently, whereas in thrillers the main conflict is external in all cases; in both genres, a stable coherence is maintained between conflict, structure, and narrative conventions. Likewise, all narratives present closed endings, suggesting a model preference for resolute and conclusive structures regardless of genre or experimental condition.

Regarding structural complexity, results show clear differences between generative models (cf. Table 3). The presence of subplots is limited to dramatic screenplays and appears systematically in texts

generated by Claude, whereas ChatGPT-produced screenplays display predominantly linear structures organized around a single main plot.

In the thriller genre, both models dispense with subplots and opt for more direct and functional narrative architectures focused on progression of the main conflict. These findings indicate that AI management of structural complexity is conditioned by narrative genre, prioritizing clarity and dramatic effectiveness over expansion of the narrative architecture.

With respect to master plots, results indicate that screenplays generated by both models align consistently with recognizable narrative typologies appropriate to genre, in line with Tobias's (1993) proposals. In dramatic screenplays, plots centered on transformation, discovery, or moral responsibility predominate, whereas in thrillers structures based on pursuit, mystery, or escape prevail. This alignment between plot type and genre remains stable across all experimental conditions, suggesting that AI models have internalized conventional narrative patterns associated with each genre and apply them coherently in structural construction.

Despite the high degree of structural adequacy across the corpus, the analysis identifies deviations and errors that primarily affect climax management and conflict resolution. In Claude-generated screenplays, anticlimactic or excessively emphasized resolutions occur more frequently, especially in thrillers and under Condition B, where the accumulation of tension does not always translate into a proportionate narrative closure.

ChatGPT-produced screenplays, by contrast, more recurrently display repetitive developments, with limited dramatic progression and little conflict evolution. Forced resolutions or solutions external to prior narrative development are also detected sporadically. Overall, these results suggest that although both models master the basic structural schemes of the cinematic screenplay, difficulties persist in the fine-grained management of dramatic progression and narrative closure as narrative complexity increases.

Results regarding the actantial definition of characters show consistent differences between the generative models by genre and experimental condition (cf. Table 4). Claude-generated screenplays systematically present clearly defined and characterized protagonists and secondary characters in both drama and thriller. By contrast, ChatGPT-produced texts tend toward reduced actantial configurations, with less developed protagonists and limited presence of secondary characters, especially in dramatic screenplays.

Table 4. Results for the *characters* variable.

	CGP CND A Drama	Claude CND A Drama	CGP CND B Drama	Claude CND B Drama	CGP CND A Thriller	Claude CND A Thriller	CGP CND B Thriller	Claude CND B Thriller
Protagonist defined and characterized	NO	YES	YES	YES	NO	YES	YES	YES
Antagonist defined and characterized	N/A	N/A	N/A	N/A	NO	YES	YES	YES
Secondary characters defined and characterized	NO	YES	NO	YES	NO	YES	YES	YES
Protagonist transformation arc	Flat	Radical	Moderate	Moderate	Flat	Flat	Flat	Flat
Antagonist transformation arc	N/A	N/A	N/A	N/A	Flat	Flat	Flat	Flat
Relational dynamics coherent and clear	NO	Sometimes	Sometimes	YES	NO	YES	NO	YES

Fuente: Own elaboration, 2026.

In thrillers, the antagonist figure appears consistently and functionally in Claude screenplays, whereas in those generated by ChatGPT antagonist characterization is more diffuse or less marked. Overall, these results indicate that Claude tends to construct more complete and hierarchical actantial systems, whereas ChatGPT prioritizes more contained and functional narrative configurations, with lower density in the definition of narrative roles.

Analysis of transformation arcs shows clear differences between the generative models depending on narrative genre (cf. Table 4). In dramatic screenplays, Claude presents protagonists with moderate or radical transformation arcs, whereas ChatGPT-generated texts tend toward flat or minimally developed arcs. In the thriller genre, both models consistently display flat transformation arcs for both protagonists and antagonists, in coherence with genre conventions. Nevertheless, Claude maintains greater functional clarity in arc definition, whereas in ChatGPT character transformation is less relevant from a structural standpoint.

Overall, results on character construction are consistent with the patterns observed in the structure variable. Screenplays that present greater structural complexity and a more expanded narrative architecture, particularly in Claude, also tend to articulate more clearly defined actantial systems and more elaborated transformation arcs. By contrast, the more contained and linear structures predominant in ChatGPT screenplays are associated with functional characters and lesser depth in character development. This correspondence reinforces the idea that AI integrates structural organization of the narrative with character construction.

Results for the space–time variable indicate a high degree of narrative control across the corpus (cf. Table 5). All generated screenplays, regardless of model, genre, or experimental condition, maintain clear organization of locations and coherent temporality, enabling a fluid and comprehensible reading of the narrative. This behavior suggests that AI models have robustly internalized basic conventions of spatial and temporal continuity typical of the cinematic screenplay, avoiding complex disruptions or broad temporal jumps that might compromise narrative clarity.

Table 5. Results for the *space* and *time* variable.

	CGP CND A Drama	Claude CND A Drama	CGP CND B Drama	Claude CND B Drama	CGP CND A Thriller	Claude CND A Thriller	CGP CND B Thriller	Claude CND B Thriller
Uses rhythm variations within scenes	NO	YES	YES	YES	NO	YES	YES	YES
Maintains clear temporality	YES	YES	YES	YES	YES	YES	YES	YES
Uses flashbacks	YES	NO	NO	NO	YES	NO	YES	YES
Uses broad temporal jumps	NO	NO	NO	NO	NO	NO	NO	NO

Fuente: Own elaboration, 2026.

In relation to rhythm and the use of temporal devices, clear differences are identified between the generative models. Claude-produced screenplays incorporate rhythm variations within scenes more consistently, whereas ChatGPT-generated texts tend toward a more homogeneous rhythm. Flashback use appears sporadically and in a controlled manner in some texts, without becoming a dominant structural device, and no broad temporal jumps are recorded in any of the analyzed screenplays, reinforcing the primacy of narrative continuity.

With respect to variables linked to the characteristics of the *drama* and *thriller* genres, results show a general adequacy of AI-generated screenplays to the narrative conventions of both drama and thriller. In dramatic screenplays, internal conflict, the weight of interpersonal relationships, and the use of dialogue as a narrative driver predominate, with transformation processes tied to emotional or moral dilemmas. These traits appear more developed in Claude-produced texts, whereas in screenplays generated by ChatGPT the treatment of drama is more functional and contained.

In thrillers, narratives are structured around external conflicts, situations of threat, and dynamics of pursuit or escape, with progression oriented toward resolving danger or mystery. Both models reproduce these patterns in recognizable ways, although Claude introduces greater narrative density and rhythmic variation, whereas ChatGPT opts for more direct and linear structures.

Analysis of *errors* in AI-generated screenplays reveals differentiated patterns by generative model and error type, as shown in Table 6. Overall, the detected errors do not compromise global intelligibility, but they do affect different levels of screenplay construction. Claude-generated screenplays more frequently exhibit surface-level and internal-coherence errors, such as isolated spelling mistakes, spatial or temporal inconsistencies, and variations in character naming, especially in thrillers and under Condition B. By contrast, ChatGPT-produced screenplays show a greater tendency toward repetition of dialogue, scenes, or situations, and toward the use of non-visual descriptions, indicating issues of redundancy and inadequate alignment with the specific language of the cinematic screenplay.

Table 6. Results for the *errors* variable.

	CGP CND A Drama	Claude CND A Drama	CGP CND B Drama	Claude CND B Drama	CGP CND A Thriller	Claude CND A Thriller	CGP CND B Thriller	Claude CND B Thriller
Spelling errors	NO	YES	NO	NO	YES	NO	YES	YES
Repetitive dialogue	YES	YES	NO	YES	NO	NO	NO	YES
Repetition of scenes or situations	YES	YES	NO	YES	NO	NO	NO	NO
Non-visual descriptions	YES	YES	NO	YES	NO	YES	YES	YES
Unnecessary flashbacks	YES	N/A	N/A	N/A	N/A	N/A	YES	YES
Character with different names	NO	NO	NO	YES	NO	NO	YES	NO
Object and/or space inconsistencies	NO	YES	NO	NO	NO	NO	YES	YES
Temporal inconsistencies	NO	YES	NO	NO	NO	NO	NO	NO

Fuente: Own elaboration, 2026.

Analysis of *biases* in AI-generated screenplays does not reveal systematic or dominant patterns across the corpus. Detected biases appear sporadically and contextually, primarily linked to character characterization and the assignment of narrative roles rather than to an explicit reproduction of rigid stereotypes. This behavior is especially evident under Condition A, where demographic neutrality instructions clearly constrain character construction and limit the explicitness of identity markers.

When *biases* do appear, they are mainly related to attributes such as age, gender, or social class, as well as to the functional assignment of certain roles within the narrative. In these cases, biases operate implicitly, embedded in widely recognizable narrative conventions rather than as features explicitly emphasized or problematized by the text itself. No consistent tendency toward extreme polarization or overtly stereotyped character construction is observed; instead, there is a moderate reproduction of conventional narrative schemas.

From a comparative perspective, subtle differences between the generative models can be observed. Claude produced screenplays that tend to verbalize certain identity related or relational features more explicitly, facilitating the identification of potential biases, whereas in ChatGPT generated texts, these emerge more implicitly and functionally, subordinated to action progression or conflict structure. Overall, the results suggest that the presence of biases in AI-generated screenplays does not follow a systematic logic, but rather emerges contextually depending on genre, experimental condition, and the narrative strategies adopted by each model.

Complementary to the external analysis of the corpus, the AI models themselves were asked to identify potential biases present in the generated screenplays. Results from this self-analysis exercise indicate that the models tend to recognize broad, general biases linked to narrative conventions and widely normalized functional roles. However, this identification is partial and imprecise, as it rarely aligns with the biases detected through the systematic analysis applied in this study. This mismatch indicates that although models can generate meta-reflexive discourse about their own texts, their capacity to identify implicit biases in narrative construction is limited and dependent on explicitly formulated categories.

4. Conclusions

The results of this study show that the artificial intelligence models analyzed are capable of generating screenplays that consistently reproduce the basic narrative conventions associated with the drama and thriller genres. Both ChatGPT and Claude internalize recognizable structural schemes, articulate genre-appropriate conflicts, and generally maintain adequate control over the spatial, temporal, and formal clarity of the narrative. These findings confirm that AI-based screenwriting cannot be understood as a random process, but rather as a practice governed by stable, learned narrative patterns.

However, the comparative analysis reveals significant differences between the models in their management of narrative complexity. Claude tends to expand the narrative universe through a greater

number of characters, locations, and structural layers, and it develops more elaborated transformation arcs for main characters in the drama genre. This greater complexity is accompanied, however, by a higher incidence of formal inconsistencies and lower stability in adhering to professional screenplay format. By contrast, ChatGPT adopts more contained and functional narrative strategies, with more linear structures and greater formal regularity, albeit with lower dramatic density and a tendency toward repetition.

Overall, the results suggest that artificial intelligence approaches screenplay creation through a logic of genre adequacy and structural control, modulated by narrative strategies that differ by model. In this regard, the contrast between the external analysis of the screenplays and the bias self-identification exercises performed by the models themselves reveals a gap between narrative production and the meta-reflexive capacity of artificial intelligence. While the models can generate general statements about potential biases in their texts, these rarely coincide with the patterns detected through systematic analysis, reinforcing the need for external analytical frameworks to critically evaluate AI-generated screenwriting. Rather than implying a replacement of human authorship, these results position artificial intelligence as a tool for creative support whose narrative and formal limitations still require specialized analytical oversight.

Referencias

- Altman, R. (2010). *Los géneros cinematográficos (2nd reimpresión)*. Espasa Libros S.L.U.
- Aumont, J., Bergala, A., Marie, M. & Vernet, M. (2019). *Estética cinematográfica: 120 años de teoría y de cine*. La Marca Editora.
- Bang, J. (2022). *Script analysis: Deconstructing screenplay fundamentals*. Routledge.
- Barthes, R. & Duisit, L. (1975). An introduction to the structural analysis of narrative. *New Literary History*, 6(2), 237-272. <https://doi.org/10.2307/468419>
- Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, 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*, 610-623. <https://doi.org/10.1145/3442188.3445922>
- Bordwell, D., Thompson, K. & Smith, J. (2020). *Film art. An introduction (12th ed.)*. McGraw-Hill Education.
- Cake, S. (2025). Artificial intelligence as a collaborative tool for script development. *Media practice and education* 26(3), 302-317. <https://doi.org/10.1080/25741136.2025.2454074>
- Carmona, R. (2016). *Cómo se comenta un texto fílmico (7th ed.)*. Cátedra.
- Casetti, F. & Di Chio, F. (2017). *Cómo analizar un film (11th reimpresión)*. Espasa.
- Chatman, S. (1980). *Story and discourse. Narrative structure in fiction and film*. Cornell University Press.
- Cruz, C. (2024). *Imágenes narradas. Cómo hacer visible lo invisible en un guion de cine (2nd ed.)*. Laertes S.A.
- Cuadrado Alvarado, A. (2017). *Narración audiovisual. Síntesis*.
- Dancyger, K. & Rush, J. (2002). *Alternative screenwriting (3rd ed.)*. Focal Press.
- Field, S. (2005). *Screenplay. The foundations of screenwriting. A step-by-step guide from concept to finished script*. Bantam Dell a Division of Random House, Inc.
- García Barrientos, J. L. (2007). *Cómo se comenta una obra de teatro (2nd reimpresión)*. Editorial Síntesis S.A.
- García Serrano, F. (2019). *Técnicas del guion: Métodos, fundamentos, estructuras y conceptos*. Síntesis.
- Manovich, L. (2020). *AI aesthetics*. Strelka Press
- Martin, M. (2008). *El lenguaje del cine*. Gedisa.
- McKee, R. (2009). *El guion. Sustancia, estructura, estilo y principios de la escritura de guiones*. Alba minus.
- McKee, R. (2022). *El personaje. El arte de crear personajes en la página, el escenario y la pantalla*. Alba Editorial S. L. U.
- Rajas, M. & Gértrudix, M. (2025). Generative Artificial Intelligence in Audiovisual Screenwriting. En A. Baraybar-Fernández, S. Arrufat-Martín y B. Díaz Díaz (ed.), *The AI Revolution. How Technological Developments Affect the Audiovisual Sector*. (pp. 91-101). Springer.
- Sánchez Escalonilla, A. (2008). *Estrategias de guion cinematográfico (6th ed.)*. Editorial Ariel, S.A.
- Schatz, T. (1981). *Hollywood genres: Formulas, filmmaking, and the studio system*. Random House Inc.
- Seger, L. (2017). *Cómo convertir un buen guion en un guion excelente (11th ed.)*. Ediciones Rialp, S.A.
- Seger, L. (2018). *El secreto del mejor cine. El subtexto en el guion y en la novela*. Ediciones Rialp, S.A.
- Tobias, R. (1993). *20 Master plots and how to build them*. Writer's Digest Books.
- Vanoye, F. (1996). *Guiones modelo y modelos de guion*. Paidós Ibérica.
- Viljoen, A. (2024). Scriptwriting with AI: A comparative analysis of human and AI-generated screenplays. *Journal of African Cinemas* 15(2&3), 199-222. https://doi.org/10.1386/jac_00101_1
- Vogler, C. (2020). *El viaje del escritor. Las estructuras míticas para escritores, guionistas, dramaturgos y novelistas (3rd ed.)*. Redbook Ediciones, S. L.
- Yorke, J. (2014). *Into the Woods. How stories work and why we tell them*. Penguin Random House Uk.
- Zimmerman, P. (2018). Los personajes: la savia de la historia. En A. Steele (Ed.), *Escribir cine. Guía práctica para guionistas de la famosa escuela de escritores de Nueva York*. Alba.