

# ***Pocket Hollywood: Transforming Film Pre-Production with AI Frameworks***

David J. Ortiz-Rivera  
Master in Computer Science  
Advisor: Dr. Nelliud Torres-Batista  
Polytechnic University of Puerto Rico  
Graduate Project EXPO, February 2025

---

**Abstract** — *Pocket Hollywood is an AI assistant designed to streamline the film pre-production process, addressing challenges in scriptwriting, storyboarding, budgeting, and logistics. Leveraging OpenAI’s generative models—GPT-4o Mini and DALL·E 3—this framework enables users to focus on their creative process by reducing the financial or technical overhead required to bring a film to life. This paper explores the framework; its inner workings, development process, content generation capabilities through prompt engineering, and the broader implications of integrating AI into filmmaking. Using “Secrets of El Morro”, we illustrate its potential to create narratives, visuals, and script breakdowns, showcasing the transformative role of AI in this industry.*

**Keywords** — *AI, AI Assistant, Film, Pre-Production.*

## **INTRODUCTION**

Film pre-production plays a significant role in filmmaking, as it helps define the initial scope of a film. It includes tasks such as scriptwriting, storyboarding, budgeting, and logistical planning [1]. For filmmakers, this phase may present significant challenges due to the upfront costs of hiring writers or artists; securing filming locations; obtaining permits and insurance; and the extensive planning required to coordinate schedules or assemble crews; before a single scene is shot. These issues can lead filmmakers to compromise on their creative vision or face delays in pre-production due to handicaps or bottlenecks.

To address these challenges, Pocket Hollywood leverages OpenAI’s generative models—GPT-4o Mini and DALL·E 3—to provide filmmakers with a tool capable of generating content such as scripts,

conceptual designs, or promotional material, alongside facilitating logistical analysis, without the need of extensive resources, specialized expertise, or large production teams.

This paper explores the framework, its inner workings, its development process, content generation capabilities through prompt engineering, and the broader implications of integrating AI into filmmaking. Through a case study of “*Secrets of El Morro*”, we illustrate its potential to create narratives, visuals, and script breakdowns. Finally, we examine the transformative role of AI in the filmmaking industry and the opportunities it offers for innovation and accessibility in the industry.

## **FRAMEWORK**

The process begins with an initial prompt, where the user provides fundamental ideas or parameters, such as the genre, setting, and key characters. This prompt is processed in the text generation stage by GPT-4o Mini, which produces textual outputs such as synopses, scripts, or logistical analysis. These outputs are then transformed and fed to the visual generation stage, where DALL·E 3 interprets text or additional user input to create conceptual or promotional visual materials.

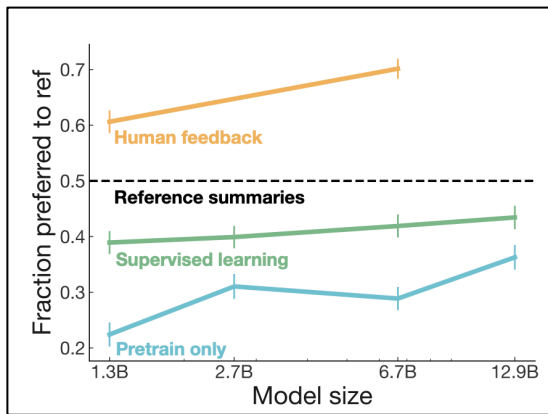
## **GPT-4O MINI / REINFORCEMENT LEARNING WITH HUMAN FEEDBACK**

Pocket Hollywood utilizes OpenAI’s GPT-4o Mini model for generating content and conducting logistical analyses. This is made possible by an advanced language model training technique called Reinforcement Learning with Human Feedback (RLHF), explored in the paper ‘*Learning to Summarize from Human Feedback*’ by N. Stiennon

et al., which focuses on improving text summary quality using language models [2].

Traditional text summarization models often relied on metrics like Recall-Oriented Understudy for Gisting Evaluation (ROUGE). This metric compares n-gram matches between model-generated summaries and human-written reference texts [3]. However, N. Stiennon et al. argue that such metrics don't necessarily correlate with summary quality or human preferences [2]. To address this, they trained a reward model using a large dataset of human-labeled comparisons between summaries, resulting in outputs more aligned with user expectations [2].

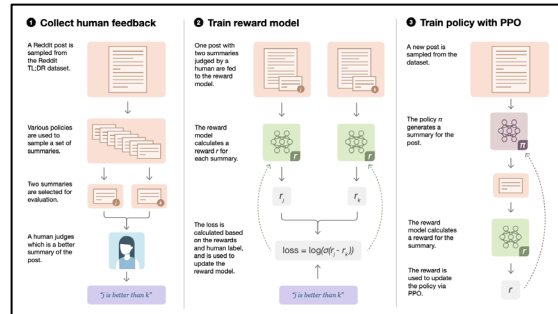
As shown in Figure 1, RLHF enables summarization models to outperform traditional methods, such as pretraining on large datasets to predict the next token or fine-tuning these pretrained models on labeled datasets. While increasing model size improves performance across all methods, RLHF demonstrates the most significant gains in alignment with human preferences. Training RLHF models larger than 6.7B parameters, however, is constrained by high computational costs, as the 6.7B model required approximately 320 GPU-days [2]. These findings underscore both the effectiveness of RLHF and the challenges of scaling it further.



**Figure 1**  
Performance of Summarization Models Trained with RLHF against Traditional Methods

Figure 2 illustrates the three key stages of the RLHF process, which centers around training and refining a policy, a model that determines how outputs are generated. Building on the results in Figure 1, this process begins with collecting

demonstration data to train a supervised policy, followed by training a reward model based on ranked outputs. Finally, the policy is optimized using Proximal Policy Optimization (PPO) [2]. For language models such as GPT-4o Mini, this process involves using prompts as the dataset, with human feedback refining the model's responses [4]. This ensures that the model generates outputs aligned with user expectations by continuously incorporating human feedback to refine its understanding of quality.



**Figure 2**  
Overview of the Reinforcement Learning with Human Feedback Process

### DALL·E 3

DALL·E 3, OpenAI's image generation model, creates visuals from textual descriptions using a diffusion-based architecture. This method iteratively refines an image from random noise, guided by input text [5]. For example, a prompt like "Give me an image of a film set at dawn" produces visuals with accurate lighting and atmospheric tones, even without explicit details. To ensure the visuals align with the user's intent, OpenAI integrates the Contrastive Language-Image Pre-training (CLIP) model, which ranks generated images based on their semantic fidelity to the prompt [6].

### DEVELOPMENT PROCESS

Initially, Pocket Hollywood was initially planned as a Python3-based script, but during setup, we found that OpenAI already provides an interface that allows users to create assistants by specifying name, system instructions, and model.

## OpenAI’s Assistant Interface

Through the assistant interface, we configured the following parameters:

- Name: Pocket Hollywood
- System Instructions: The assistant was given a clear role as a creative AI specializing in film. The instructions emphasize professional yet collaborative behavior, a creative tone, and flexibility in adapting to user preferences (see Table 1 for full system instructions).
- Model: GPT-4o Mini
- Response Format: Text.
- Temperature and Top-P: Both were left with their default values to avoid unwanted randomness or lack of coherence.

### System Instructions

These instructions, as detailed in Table 1, established Pocket Hollywood’s tone, creative scope, and ethical considerations:

**Table 1**  
System Instructions for Pocket Hollywood

Aspect	Instructions
Primary Goal	Assist filmmakers, writers, and enthusiasts by generating creative content.
Tone and Style	Professional, friendly, adaptable to genre and style.
Creative Outputs	Generate scripts, scenes, storyboards, and alternative ideas.
Cinematic Knowledge	Reference filmmaking techniques, story arcs, and character development.
Ethics and Sensitivity	Avoid controversial content and ensure diversity and inclusivity.
User Collaboration	Encourage collaboration through questions and feedback integration.

### Considerations on Prompt Engineering

Prompt engineering is crucial to maximize a language model’s capabilities effectively, as the phrasing and structure of prompts can make the help align its intent and output. Well-crafted prompts guide the model to understand the user’s intent clearly, reducing ambiguity in the process. For example, asking the model to “Describe a dramatic

scene where a detective confronts a criminal” might yield a generic or overly broad response. In contrast, a more detailed prompt such as “Write a dialogue-driven scene where a detective confronts a criminal in a dimly lit alley during a thunderstorm, focusing on tension and power dynamics”, provides the model with context, setting, and tone. This specificity not only enhances the relevance of the generated content, but also allows the assistant to showcase creativity within a defined framework.

## GENERATED CONTENT

In this section, we present a series of prompts alongside the generated content, showcasing the assistant’s capabilities in text and image creation.

### Synopsis

#### Prompt:

“Write a synopsis for a movie based in El Morro that takes inspiration in The Davinci Code.”

#### Output:

##### *Secrets of El Morro*

*In the vibrant landscapes of modern Puerto Rico, renowned historian and cryptologist Dr. Clara Vega finds herself drawn into a web of intrigue when she stumbles upon a centuries-old diary hidden deep within the archives of San Juan. The diary, once belonging to one of the island’s 16th-century Spanish governors, hints at the existence of a secret treasure linked to the legendary fortress of El Morro. According to local lore, this treasure is said to hold not just material wealth but also powerful artifacts that could alter the course of history.*

*As news of the discovery leaks, a nefarious secret society, the Caballeros del Mar, becomes obsessed with acquiring the knowledge embedded in the diary. Led by the enigmatic and ruthless Rafael Luque, they will stop at nothing to uncover the treasure for their own gain. Unbeknownst to Clara, she is also connected to the treasure through her family lineage—her ancestors were once guardians of the fortress, entrusted with protecting its secrets.*

*With the assistance of Javier, a charismatic local archaeologist harboring secrets of his own,*

*Clara races against time as the Caballeros del Mar closes in. Their adventure takes them through the ancient tunnels of El Morro, the labyrinthine streets of Old San Juan, and hidden coastal ruins, piecing together clues left by history's guardians. Along the way, Clara uncovers stunning revelations about her heritage, the true significance of El Morro, and its role in a hidden chapter of the island's past.*

*In a breathtaking climax filled with twists, coded messages, and exhilarating action, Clara realizes that the treasure's true power lies in uniting the people of Puerto Rico rather than in wealth—sparking a movement that celebrates the island's rich history and resilience. As Clara confronts Luque in a final showdown amidst the crashing waves at El Morro, she must decide between personal glory and the greater good of her homeland.*

*As the sun sets over the storied fortress, the film leaves audiences questioning how history shapes our identities and the strength found in uncovering and embracing the secrets of the past.*

### **Scene / Script Breakdown**

#### **Prompt:**

“Based on this synopsis, give me a scene from the movie along with a script breakdown for props, locations, characters, etc.”

#### **Content:**

*Scene: The Discovery of the Diary*

*Setting: A dimly lit archive room in San Juan, filled with dusty shelves lined with ancient books and manuscripts. The walls are adorned with old maps of Puerto Rico and relics from the Spanish era. Shafts of sunlight filter in through small, barred windows, casting an eerie glow on the artifacts.*

*Script Breakdown*

*Characters:*

*Dr. Clara Vega - A passionate historian in her late 30s. She is determined, intelligent, and possesses a deep love for her heritage.*

*Javier - A charismatic local archaeologist in his early 40s, charming yet mysterious. He has a deep knowledge of Puerto Rican history and shares a hidden connection to Clara.*

*Luis - Clara's assistant, a tech-savvy college student in his 20s, providing comic relief and enthusiasm, balancing Clara's seriousness.*

*Props:*

*An old, leather-bound diary with ornate clasps and yellowed pages.*

*A magnifying glass for examining entries.*

*A lantern to light in the dark, dusty archives.*

*An ancient map of Old San Juan with marked locations.*

*A laptop for Luis, connecting him to online archives and databases.*

*Locations:*

*Location 1: Archive Room (Indoor)*

*Location 2: Nearby Plaza (Outdoor), where the characters briefly discuss their findings.*

*INT. ARCHIVE ROOM - DAY*

*The camera pans through the dusty archive, revealing a treasure trove of historical documents. CLARA leans over a cluttered wooden table, her brow furrowed in concentration as she flips through a stack of old manuscripts.*

*CLARA (whispers to herself) Come on... there has to be something here.*

*LUIS enters, energetic and clutching his laptop.*

*LUIS (excited) Dr. Vega! I found a digital manuscript about El Morro from the 1600s. It might contain references to—*

*He suddenly stops, eyes wide, noticing the diary clasped in her hands.*

*CLARA (looking up, startled) Luis! Look at this!*

*She opens the diary, and the camera zooms in on the ornate writing, revealing sketches of the fortress and cryptic symbols.*

*JAVIER (O.S.) (entering casually) What have you uncovered?*

*JAVIER walks in, leaning against the door frame, a playful smile on his face. CLARA looks up, her voice a mix of excitement and urgency.*

*CLARA (pointing to the diary) This is a governor's diary from the 16th century. It talks about a treasure tied to El Morro.*

*JAVIER approaches, eyes narrowing in curiosity as he inspects the pages.*

*JAVIER (challenging) And you believe it's real?*

Suddenly, the sound of distant voices echoes down the corridor, causing all three to freeze. A sense of urgency fills the air.

LUIS (nervously) Do you think someone else knows about this?

CLARA, her expression determined, looks between her two companions.

CLARA We have to be careful. The Caballeros del Mar might not be far behind.

As they huddle closer around the diary, the camera zooms out to reveal the vast archive filled with history—hinting at the adventure that lies ahead.

END SCENE

Next Steps:

Location Scout: Identify a suitable archive room, preferably in a historic building in San Juan.

Casting: Cast actors that can embody the characters' complexities.

Props Acquisition: Create or source the diary, maps, and other elements that enhance the authenticity of the setting.

Cinematography: Plan for varied camera angles to capture the intensity of discovery and the vibe of hidden histories.

### Poster

#### Prompt:

“A cinematic poster for the action-adventure thriller movie titled ‘Secrets of El Morro’. The background features the iconic El Morro fortress in Puerto Rico at sunset, with golden and orange tones blending with stormy clouds in the distance. In the foreground, the main character, Dr. Clara Vega, a determined woman with long hair blowing in the wind, holds an ancient diary. Behind her stands Javier, a charismatic archaeologist holding a map and archaeological tools. To their side, a shadowy figure of Rafael Luque, the villain, looms ominously. Symbols of ancient carvings and a faint map of Puerto Rico overlay the scene. The tagline reads: ‘A hidden legacy. An ancient mystery. The fate of the island is in their hands.’ The design is dramatic and action-packed, with a sense of mystery and historical intrigue.”



Figure 3  
Secrets of El Morro Movie Poster

### DISCUSSION

The functionality of Pocket Hollywood highlights both its potential and limitations, which impact the overall workflow and user experience of the framework. One challenge is the separation between text and visual models, which operate as independent entities. This setup requires additional steps to create visuals, taxing the user with effort. For instance, while Pocket Hollywood generates detailed prompts for DALL·E 3, users must manually transfer these prompts to the visual model to produce images. To mitigate this, we include instructions in prompts for Pocket Hollywood to generate highly detailed text-to-image-based prompts tailored for direct use in DALL·E 3. However, an ideal solution would involve a unified platform that integrates text and visual outputs, enabling more fluid and efficient collaboration between these components.

Another challenge involves inconsistencies in the accuracy and quality of the images produced by DALL·E 3, even with the use of detailed prompts. For example, the generated poster erroneously has an out of place text element between “A Hidden Legacy” and “An Ancient Mystery”, underscoring the limitation of visual generative models capable and the importance of iterative refinement.

The increasing adoption of AI in the film industry, introduces some ethical considerations. How does one define originality in the context of

content created by algorithms trained on large amounts of publicly available datasets? Ownership is another significant dilemma: who retains the rights to AI-generated visuals—the creator of the prompt, the tool’s developers, or the dataset contributors? Could a filmmaker unknowingly produce a storyboard or concept art derivative of an existing copyrighted work? If so, how do these risks scale for commercial productions? How can we prevent studios from leveraging AI as a cost-cutting tool at the expense of creative integrity and human employment? Why does the integration of AI in artistic domains often face skepticism, with critics arguing that it diminishes the authenticity and emotional depth that only human creators can provide? As AI systems like Pocket Hollywood become increasingly integrated into creative workflows, addressing these concerns is essential to preserve the integrity of the arts, ensure ethical compliance, and mitigate potential legal disputes.

While these limitations and questions are present, this framework ultimately provides a foundation for prototyping during the process of film pre-production, making it an accessible tool for established or aspiring filmmakers to explore ideas without the financial or technical overhead of traditional methods. The synopsis, scene and script breakdown for “*Secrets of El Morro*” effectively captured the intent of the respective prompts. Its ability to generate detailed script breakdowns—including props, locations, and characters—offers benefits for budget scoping and the overall logistical planning.

## CONCLUSION

Pocket Hollywood demonstrates the transformative role of AI in the filmmaking industry by enabling the intent-aligned generation of narratives, visuals, and script breakdowns, as illustrated in the case study of “*Secrets of El Morro*”. By leveraging OpenAI’s generative models—enhanced through techniques like Reinforcement Learning with Human Feedback—the framework

manages to reduce financial and technical overhead required to bring films to life.

However, this isn’t without its flaws. The current separation between text and visual models, taxes the user with additional effort required to produce desired visuals. On the other hand, inconsistencies in image quality and accuracy underscores the limitations of visual generative models and its greater need for iterative refinement. Furthermore, ethical concerns surrounding originality, ownership, and representation must be addressed to ensure responsible adoption of these technologies in creative workflows. These challenges, also represent opportunities for future improvements, such as developing unified platforms and refining model capabilities to better align outputs with intent.

Despite these limitations, Pocket Hollywood offers a foundation for prototyping, helping filmmakers conceptualize ideas more efficiently as it generally aligns with the prompt crafted by the user. Continued advancements in AI and model refinement present opportunities to address these limitations, paving the way for more cohesive, accurate, and user-friendly tools that support creative processes without compromising artistic intent.

## REFERENCES

- [1] Studio Binder. (2024, December 1). *What is Pre-Production? The Complete Guide* [Online]. Available: <https://www.studiobinder.com/blog/what-is-pre-production-definition>.
- [2] N. Stiennon et al. (2020, September 4). *Learning to summarize with human feedback* [Online]. Available: <https://arxiv.org/abs/2009.01325>.
- [3] C. Y. Lin. (2004, July 1). *ROUGE: A Package for Automatic Evaluation of Summaries* [Online]. Available: <https://aclanthology.org/W04-1013>.
- [4] OpenAI. (2022, November 30). *Introducing ChatGPT* [Online]. Available: <https://openai.com/index/chatgpt>.
- [5] OpenAI. (2023, October 3). *DALL·E 3 System Card* [Online]. Available: [https://cdn.openai.com/papers/DALL\\_E\\_3\\_System\\_Card.pdf](https://cdn.openai.com/papers/DALL_E_3_System_Card.pdf).

- [6] A. Ramesh, P. Dhariwal, A. Nichol, C. Chu, y M. Chen. (2022, April 13). *Hierarchical Text-Conditional Image Generation with CLIP Latents* [Online]. Available: <https://arxiv.org/abs/2204.06125>.