



Author: David J. Ortiz-Rivera
 Advisor: Dr. Nelliud Torres-Batista
 Computer Science

Abstract

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 by defining a film’s initial scope through tasks such as scriptwriting, storyboarding, budgeting, and logistical planning [1]. Filmmakers often face challenges due to upfront costs of hiring writers or artists, securing locations, obtaining permits, and coordinating schedules, which can lead to compromises in creative vision or delays. To address these issues, *Pocket Hollywood* leverages OpenAI’s generative models—GPT-4o Mini and DALL-E 3—to generate scripts, conceptual designs, and promotional material while facilitating logistical analysis without requiring extensive resources, specialized expertise, or large production teams. This paper explores the framework, development process, and content generation capabilities through prompt engineering, along with the broader implications of AI in filmmaking. A case study of *Secrets of El Morro* illustrates its potential for narrative creation, visuals, and script breakdowns, highlighting AI’s transformative role in fostering innovation and accessibility in the industry.

Background

Pocket Hollywood automates film pre-production by combining GPT-4o Mini for text generation and DALL-E 3 for visual creation. Users provide a prompt specifying genre, setting, and key elements, which GPT-4o Mini processes to generate synopses, scripts, and logistical analyses. These outputs then guide DALL-E 3, which interprets the text to create concept art and promotional visuals. GPT-4o Mini leverages Reinforcement Learning with Human Feedback (RLHF) to refine its responses, ensuring higher alignment with user expectations [2]. This technique, introduced by Stiennon et al. (2020), improves model performance by training on human-labeled comparisons [2]. However, RLHF models require significant computational resources, posing scalability challenges [2].

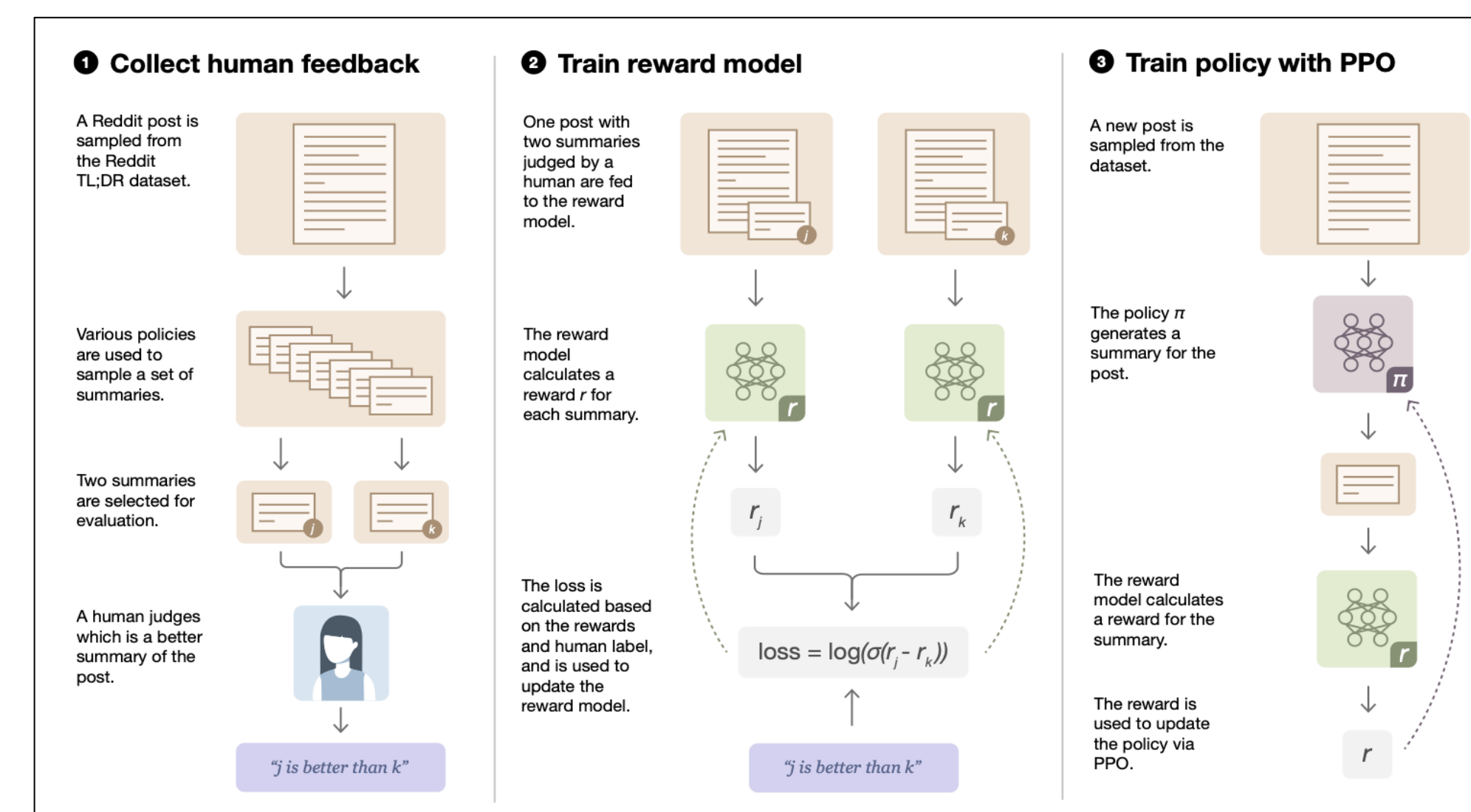


Figure 2

Overview of the Reinforcement Learning with Human Feedback Process

The RLHF 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].

Problem

The main goal of this assistant is to apply language models to real-world scenarios that help standardize processes, allowing other aspects of filmmaking to take center stage. Film pre-production is a complex and resource-intensive phase, involving scriptwriting, storyboarding, budgeting, and logistical planning, all of which can create financial and technical barriers for filmmakers. To address these challenges, *Pocket Hollywood* leverages GPT-4o Mini and DALL-E 3 to automate key aspects of pre-production, generating scripts, visuals, and production plans. By streamlining these foundational tasks, *Pocket Hollywood* aims to reduce costs, enhance accessibility, and allow filmmakers to focus more on creative execution.

Methodology

Pocket Hollywood was built using OpenAI’s Assistant Interface, which allowed configuration of AI assistant by specifying parameters like name, language model, and system instructions such as goal, tone and style. For cases requiring image based content, the assistant was tasked with generating ideal prompts to achieve higher quality outputs. 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. This provided specificity has the ability to enhance the relevance of the generated content, while allowing the assistant to showcase creativity within a defined framework.

Results and Discussion

Synopsis Prompt:

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

Output:

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,

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.”

Output:

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.

Props:

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

A magnifying glass for examining entries.

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, ...”

Output:



The framework showcases both its potential and limitations in film pre-production. A key challenge is the separation of text and visual models, requiring users to manually transfer prompts for image generation, reducing workflow efficiency. While detailed text-to-image prompts help mitigate this, an integrated platform would enhance usability. Additionally, image inconsistencies highlight the need for iterative refinement. We can see out of place text elements in the generated poster.

The increasing adoption of AI raises ethical concerns regarding originality, ownership, and AI’s impact on human creators. Despite these challenges, *Pocket Hollywood* provides a valuable foundation for prototyping, enabling filmmakers to explore ideas efficiently while reducing financial and technical barriers.

Conclusions

Pocket Hollywood demonstrates AI’s potential in filmmaking by generating narratives, visuals, and script breakdowns, reducing financial and technical barriers in pre-production, as illustrated using *Secrets of El Morro*. Leveraging OpenAI’s models with Reinforcement Learning with Human Feedback, it enhances content generation but not without its limitations. The separation of text and visual models requires manual effort, while image inconsistencies highlight the need for refinement and prompt engineering. Ethical concerns regarding originality, ownership, and representation must also be addressed. Despite these challenges, *Pocket Hollywood* provides a strong foundation for AI-assisted prototyping.

Future Work

Future work will focus on integrating *Pocket Hollywood* into a unified platform that combines text and visual generation, reducing the need for manual input transfers. Improvements in prompt engineering will be explored to enhance AI-generated content accuracy and consistency. Additionally, addressing ethical concerns regarding originality, ownership, and AI bias will be a priority to ensure responsible adoption in filmmaking. Further research will also investigate the scalability of AI-driven pre-production tools for larger, industry-level productions.

Acknowledgements

I thank Daimarik Torres Cruz for her editorial support and Dr. Nelliud Torres-Batista for his mentorship and guidance throughout this project.

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.