



Carnegie Mellon University

Story Generation

Generative AI Lab - Final Project - Group 9

Gitansh Wadhwa, Skand Vijay, Yash Thakkar

Preface

Code vs No-Code

Project Prompt Interpretations

Project Overview

Objective: Explore how LLMs and diffusion models can work together to analyze text and generate new stories and meaningful story illustrations.

Approach:

1. **Define Topic/Theme** – Select relevant themes or story title.
2. **Web Crawling** – Fetch related storybook content.
3. **Feature Extraction** – Summarize & extract key story elements.
4. **Story Generation** – Create new stories based on extracted features.
5. **Illustration Generation** – Convert scenes into AI-generated images.
6. **Comparison & Analysis** – Evaluate AI vs. existing storybooks.

Story Theme Selection

Fantasy

Stories involving magic or imaginary worlds

History

Tales focused on exciting journeys, quests, or challenges

Horror

Stories featuring suspenseful, scary, or supernatural elements designed to evoke feelings of fear or unease

Methodologies and Task Overview

Data Acquisition & Processing

- Implemented Firescraper for efficient web-scraping
- Utilized OpenAI GPT-4.0 API to extract most relevant features from online storybooks

Story Generation

- Utilized OpenAI GPT-4.0 API to generate stories based on the extracted features

Image Generation Pipeline

- Integrated DALL-E 3 API with extracted feature sets
- Systematically refined prompts through iterative output evaluation
- Replicated methodology with Stable Diffusion 3 Turbo

Comparative Analysis

- Performed qualitative assessment of image models

System Integration

- Developed comprehensive agentic workflow architecture

Story - 1: How to Train your Dragon (Fantasy)



Key features extracted

Setting: Viking Age on the island of Berk

Hiccup Horrendous Haddock III

Appearance: Slim build, messy hair, prosthetic left foot

Attire: Rustic Viking clothing with fur vest

Personality: Inventive, kind-hearted, empathetic

Toothless

Appearance: Black scales, retractable teeth, large eyes

Distinguishing Features: Rare Night Fury dragon

Personality: Playful, loyal, somewhat insecure

Environmental Design

Location: Rugged island "north of Hopeless"

Architecture: Sturdy wooden Viking structures

Landscape Features: Misty mountains, harbor

Comparing Narratives

Aspect	Actual Story (How to Train Your Dragon)	Generated Story (How to Forgive Your Dragon)
Conflict Shift	Focus on training dragons and proving one's worth	Internal struggles of forgiveness and unity
Themes	Bravery, Intelligence and defying expectations	Emotional intelligence, reconciliation, and balancing tradition with progress.
Worldbuilding	Focused on Viking society and their evolving relationship with dragons.	Introduces a prophecy, a hidden dragon council, and internal struggles within dragon society.
Tone and Style Change	Lighthearted, humorous, and action-driven.	Mature, introspective, and poetic, with themes of destiny and legacy.

Comparing Illustrations

Original Image



DALL-E 3 Output



Stable Diffusion Output



Findings & Analysis

Aspect	DALL-E 3	Stable Diffusion 3.3 Turbo
Style and Composition	Cinematic, photorealistic style with dramatic sunset lighting	Stylized, illustrated approach with children's book aesthetic
Feature Preservation	Maintained core theme of human-dragon relationship	Significantly transformed the aesthetic toward children's illustration
Technical Quality	Consistent photorealistic rendering throughout	Effective use of negative space and simplified forms
Storytelling Elements	Suggests high-stakes adventure with additional dragons in background	Implies a society living alongside dragons but with some distance



Story - 2: The Shining (Horror)

Key features extracted

Setting: Winter of 1980, Overlook Hotel, Colorado

Danny Torrance

Appearance: Short stature, innocent face, expressive eyes

Attire: Child's winter clothing, colorful jacket

Personality: Intuitive, frightened, occasionally brave

Age: 5 years old



Atmospheric Sequence

Danny riding tricycle through empty corridors (muted tones with stark light/dark contrasts), Increasing claustrophobia, psychological tension, and supernatural dread

Environmental Design

Location: Isolated Overlook Hotel in the Colorado mountains

Key Environments: Ornate hotel interiors, hedge maze

Atmosphere: Increasingly claustrophobic and threatening

Comparing Narratives

Aspect	Actual Story (The Shining)	Generated Story (Echoes of the Shining)
Conflict Shift	Jack succumbs to the hotel's influence, turning violent. The Overlook works through him. Sees past	The hotel directly targets Danny, manipulating his <i>shining</i> rather than using Jack as a vessel. Sees future
Themes	Addiction, isolation, abuse, psychological horror.	Supernatural horror, psychological manipulation, Danny's power as a gateway to the hotel's past
Worldbuilding	Overlook as a sentient antagonist with a dark history. The hedge maze is minimal (book focus is on topiary animals)	The hedge maze is central. The Overlook feels more like a cursed force preying on Danny
Tone and Style Change	Slow-burning psychological terror, deeply personal struggle	More cinematic, heightened suspense, and direct supernatural horror.



Comparing Illustrations

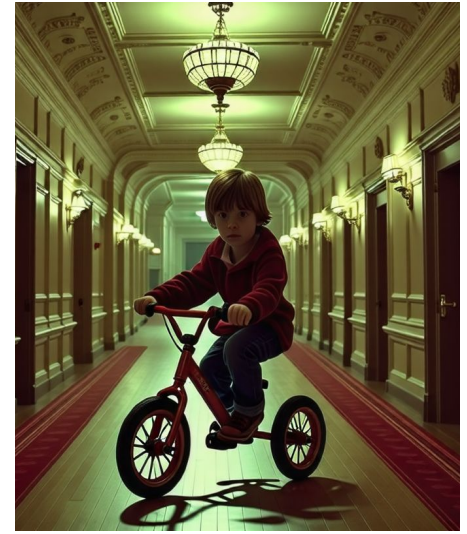
Original Image



DALL- E 3 Output



Stable Diffusion Output



Findings and Analysis

Aspect	Original Image	DALL-E 3	Stable Diffusion 3.3 Turbo
Style and Composition	Cinematic, polished, warm lighting	Animated, cool lighting	Dark, eerie, horror-like tones
Feature Preservation	Have tricycle, iconic setting, and original scene composition	Maintains tricycle & setting but alters details	Changes elements (e.g., bicycle instead of tricycle)
Technical Quality	High detail, realistic textures, film-like finish	High detail and contrast	Stylized, slightly less refined
Storytelling Elements	Mysterious but balanced, subtle unease	Mysterious but balanced	More intense, unsettling atmosphere

Story - 3: Napoleon Bonaparte (History)

Key features extracted

Setting: Corsican hills, urban Paris, European battlefields

Napoleon Bonaparte

Appearance: Thin build, average height for his era, close-cropped haircut

Attire: Military uniform or imperial robes

Personality: Ambitious, intelligent, hardworking, commanding

Visual Design

Dark with bright uniform accents

Dynamic angles

Themes

Ambition and Power

Revolution and Reform

Nationalism and Identity



Comparing Narratives

Aspect	Actual Story (War with Britain)	Generated Story (The Fall of the Empire)
Conflict	Primarily external: Napoleon vs. European powers	Both external and internal: Napoleon vs. European powers and his own conscience
Themes	Conquest, military strategy, political maneuvering	Ambition, power, revolution, nationalism, fall from grace
Worldbuilding	Factual, historically accurate depiction of 19th century Europe	More atmospheric, emphasizing emotional weight of locations
Tone	Objective, historical	Introspective, philosophical



Comparing Illustrations

Original Image



DALL- E 3 Output



Stable Diffusion Output



Findings and Analysis

Aspect	Original Image	DALL- E 3	Stable Diffusion
Style & Composition	Classical, realistic, historical painting	Dramatic, cinematic, high contrast	Stylized, vibrant, slightly surreal
Feature Preservation	Accurate depiction of Napoleon and the Russia Battle	Missed the burning building	Missed the burning buildings
Technical Quality	Detailed textures, depth, and lighting	High resolution with sharp contrasts	Smooth gradients, bold colors, artistic rendering
Storytelling Elements	Focus on historical accuracy and events	Highlights intensity and chaos of battle	Highlights intensity and chaos of battle

Agentic workflow

1 Enter Title/Theme

2 Tavily Search

3 Extract URL

4 Scrape

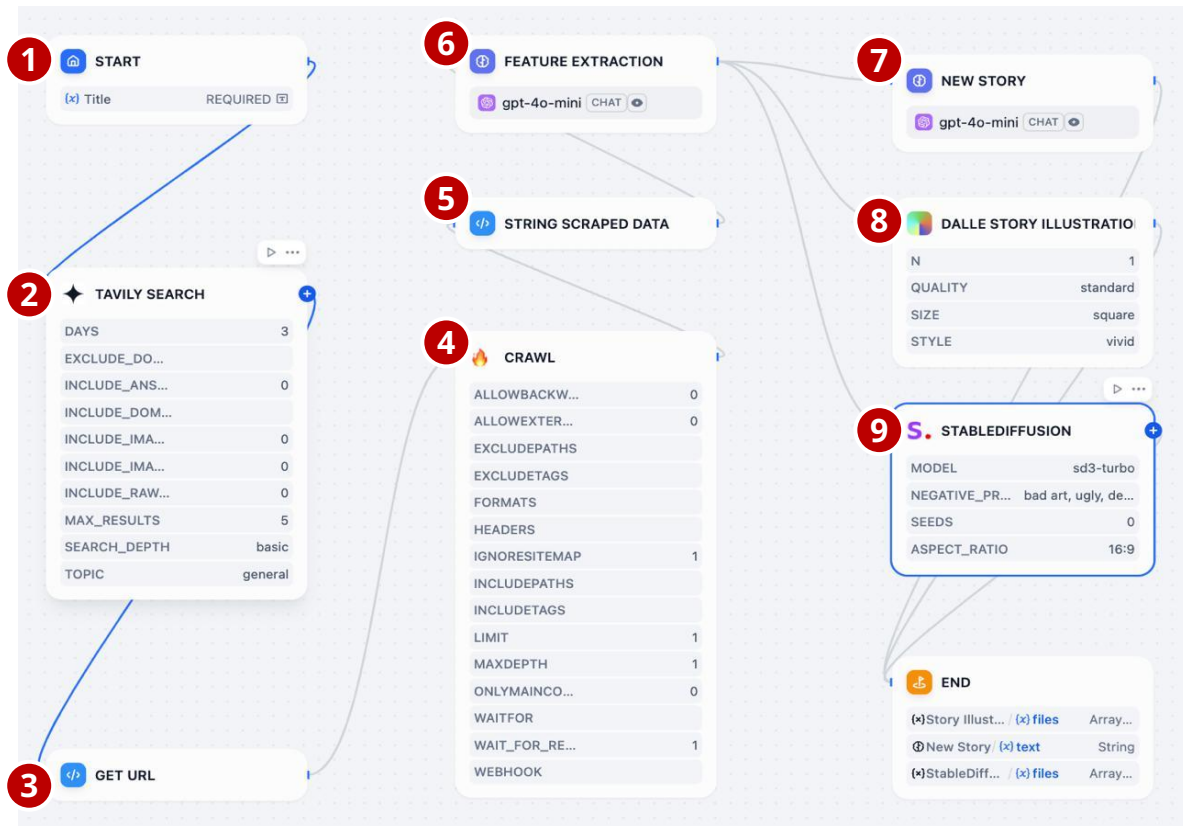
5 Get Refined Data

6 Extract Key Features

7 Generate New Story

8 Illustration Generation with DALL-E


9 Illustration Generation with Stable Diffusion



Agentic workflow - Feature Extraction


JSON

- Title
- Author
- Characters
 - Name
 - Description
 - Age
 - Clothing
 - Physical Traits
 - Personality
- Setting
 - Time Period
 - Location
 - Visual Details
- Plot Summary
- Themes
- Illustration Cues
 - Key Scenes → Scene Description
 - Mood
 - Color Palette
 - Perspective

 **Feature Extraction** ▶ 📖 ... ✕

Add description...

MODEL

 gpt-4o-mini CHAT 👁 ⌵

CONTEXT ?

(x) Crawl / json Array[Object]

To enable the context feature, please fill in the context variable in PROMPT.

SYSTEM ? 804 ★ Jinja 🔇 (x) 📋 ↗

You are an AI feature extractor specialized in breaking down stories into key elements for content generation and illustration.

Given a book summary, extract structured story details, ensuring they include visual cues for illustration.

Prompt- Generating Illustrations

S. StableDiffusion ▶ □ ... | ✕

Add description...

INPUT VARIABLES

Prompt String **Required**

Stable diffusion -

You are an expert AI artist creating high-quality book illustrations.
Generate an image based on the provided scene description.
Ensure the illustration follows these style details:
④ Feature ... / (x) text

The illustration should be visually appealing, consistent with the story's tone.

used for generating images

MODEL * ?

Stable Diffusion 3 Turbo ✕

NEGATIVE PROMPT ?

bad art, ugly, deformed, watermark, duplicated, disconti

Dalle Story Illustrations ▶ □ ... | ✕

Add description...

INPUT VARIABLES

Prompt String **Required**

dalle -

Generate a highly detailed and cinematic digital image inspired by the following JSON data. The scene should visually depict the described characters, settings, and key moments with a rich color palette, dynamic lighting, and a dramatic or immersive composition. Ensure historical or fantasy accuracy based on the provided time period and clothing styles. Use the listed moods and perspectives to enhance the depth and storytelling within the image.

④ Feature ... / (x) text

Image prompt, you can check the official documentation of DALL·E 3

IMAGE SIZE * ?

Square(1024x1024) ✕

Prompt- New Story Generation

MODEL

 gpt-4o-mini

CHAT










CONTEXT ?


{x} Set variable

SYSTEM ?

678  | Jinja ☐ {x}  

Using the following structured data as inspiration, generate a compelling new story that maintains the core essence of the characters, setting, and themes while introducing fresh perspectives, alternate plotlines, or imaginative reinterpretations. The story should be engaging, well-structured, and creative, exploring unique conflicts, relationships, and world-building elements.

Data:

 Feature ... / {x} text

Ensure the new story maintains thematic depth, dynamic character interactions, and immersive storytelling. Optionally, you may introduce speculative elements, historical deviations, or genre reimaginings while staying true to the original spirit of the narrative.

Platform Challenges

Custom Code
Output Limit
(8,000
Characters)

Multi-Agent
Workflow
Support

Error
Handling

Errors due to
frequent
version updates

Workflow Challenges

Prompt
Engineering

Generating
valid JSON
from LLM

Scraping result
(Videos)

Learnings

1

AI Models – Explored how diffusion models generate images based on extracted textual descriptions

2

LLM Output Structuring – Ensuring LLM generates valid JSON requires prompt tuning and post-processing to handle inconsistencies

3

Image-Text Alignment – Stable Diffusion sometimes misinterprets scene descriptions, requiring iterative refinement

4

Workflow Creation – Understood how to connect LLM and diffusion models into a seamless workflow

Business Use Cases

\$221B+

Gaming & Interactive
fiction industry.

\$24B+

eBook and
self-publishing industry

\$0.67B+

Personalized children's
book market

Sources

[Business Research Insights](#), [Future Market Insights](#), [Spherical Insights](#)

Q&A