Integrating Open AI into An R Shiny App: A Tool to Generate Pictures from Stories

Yunlei Tang, Danhe Tang, Prof. Matthew Lanham
Purdue University, Daniels School of Business
tang632@purdue.edu, tang631@purdue.edu, lanhamm@purdue.edu

ABSTRACT

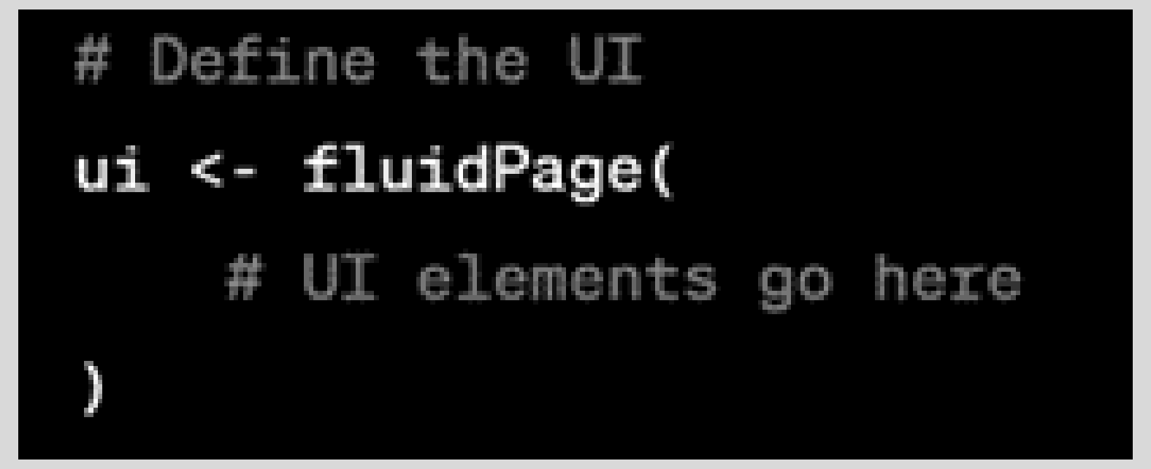
We develop and test a 'Story Picture Generator' application using Open Artificial Intelligence (AI). The tool provides the user a R Shiny Graphical User Interface (GUI), that easily interfaces with OpenAI which provides text comprehension and drawing capabilities. By leveraging OpenAI's capabilities, the application generates relevant images based on the context and themes of the user's story. The motivation for this research is that there exist diverse audiences who seek custom images to support their stories and presentations. Presenters are often tasked with trying to source something "close" to what they are seeking by searching the internet, which is often not exactly what they want, and can be time-consuming. By using this picture generator, users can visualize their text content to amplify the message, aid in comprehension, and make the content more memorable. The Open AI Story Picture Generator aims to bridge this gap, offering an automated solution to provide visual augmentations tailored to the narrative's context. The application utilizes an R Shiny interactive interface where users can input their textual content to generate a corresponding custom image. By submitting the content, the app processes the text, extracting keywords and themes by connecting and using OpenAI's Large Language Models (LLMs). These are then sent to drawing OpenAI to return appropriate images. The Shiny app then displays these images alongside the story. The app is built to be intuitive, requiring no technical expertise from the user, but offers options for image customization and refinement as necessary

INTRODUCTION

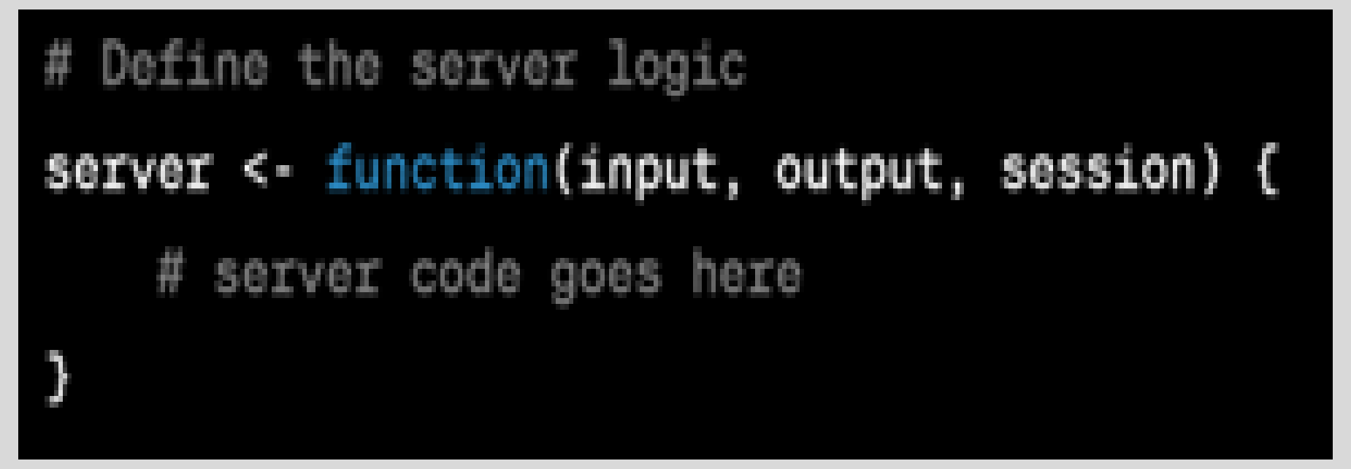
What is Shiny

Shiny is package that makes it easy to build interactive desktop and web applications straight from R & Python. R Shiny harnesses the computational, statistical, and graphical power of R and channels it into interactive, web-based dashboards. It's a transformative tool that has reshaped the way data-driven decisions are made in many fields.

- **Shiny Dashboarding**: Shiny allows users to build interactive dashboards, using for example, widgets like sliders, buttons, and dropdown menus to allow users to manipulate data in real-time and see results immediately. Outputs on Shiny dashboard, like plots or tables, are automatically updated when user inputs change. There are two key components of the R shiny:
- **UI:** The user interface (UI), where analysts specify what they want their app looking like. This typically includes widgets like sliders, buttons, and plots.
- **Server**: The server function contains the instructions that your computer needs to build the app. This is where developers tell the app what to do when a user moves a slider or clicks a button. It's basically the logic behind the app.



```
# Define the UI
ui <- fluidPage(
    # UI elements go here
)
```



```
# Define the server logic
server <- function(input, output, session) {
    # server code goes here
}
```

What is OpenAI

Founded in 2015, OpenAI artificial intelligence (https://openai.com/) research and development company focus on researching and developing AI-based tools. OpenAI is known for its research in various areas of AI, including deep learning, reinforcement learning, and robotics. The organization has been at the forefront of developing large language models (LLMs), the GPT series like GPT-3, GPT-4 which we are familiar, are results of the development to LLMs, and other high-profile AI models like DALL·E. People are taking advantage of these AI-based tools to realize personal goal by inputting simple words, or source the Open AI to create personalized AI tools.

How can these be used together?

Given Shiny's robust interactive features and OpenAI's LLMs' precise text comprehension, we aim to combine R Shiny with OpenAI to develop innovative and useful tools. In this project, by combining Shiny's interactive capabilities with DALL·E's powerful image generation (driven by OpenAI), we can create a tool that transforms textual narratives into visual images, making storytelling more engaging and accessible. This is useful because the user does not need to know or use code or have other technical knowledge to use it for their work.

RESEARCH QUESTIONS

RQ1: How accurately and effectively does the AI interpret and translate textual content into relevant and contextually appropriate images?
RQ2: What are the limitations and challenges in ensuring the generated images align closely with the themes and messages of stories?

LITERATURE REVIEW

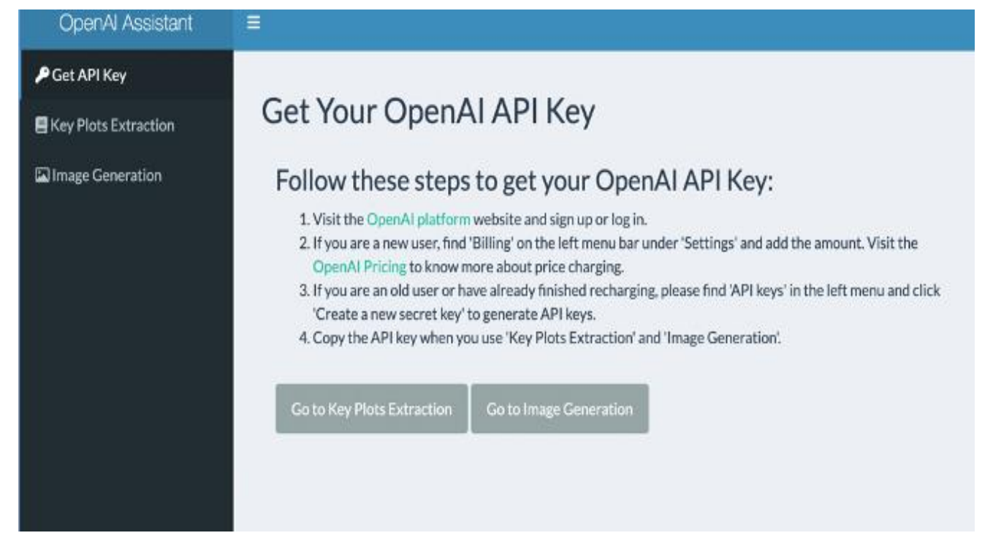
Our study focuses on the use of GPT 4.0 to capture story keywords and DALL·E 3 to generate the customized images. We found some studies invested these AI technologies in isolation or considered others.

Author, Year	Article	GPT 4.0	DALLE-3	MidJourney	DALLE-2	Craiyon
Oppenlaender, J., 2022	The creativity of text-to-image generation	No	No	**Yes**	**Yes**	No
Parsons, 2022	The DALLE-2 Prompt Book	No	No	No	**Yes**	No
Reviriego, P., & Merino-Gómez, E., 2022	Text to image generation: leaving no language behind	No	**Yes**	**Yes**	No	**Yes**
Zeqiang, L. et. al, 2023	Mini-DALLE3: Interactive Text to Image by Prompting Large Language Models	No	**Yes**	No	No	No
Our Study		**Yes**	**Yes**	No	No	No

Table 1: Literature Review Comparison of Technologies Used to Generate Custom Images

SHINY APP

The panel of the application is shown as follow:



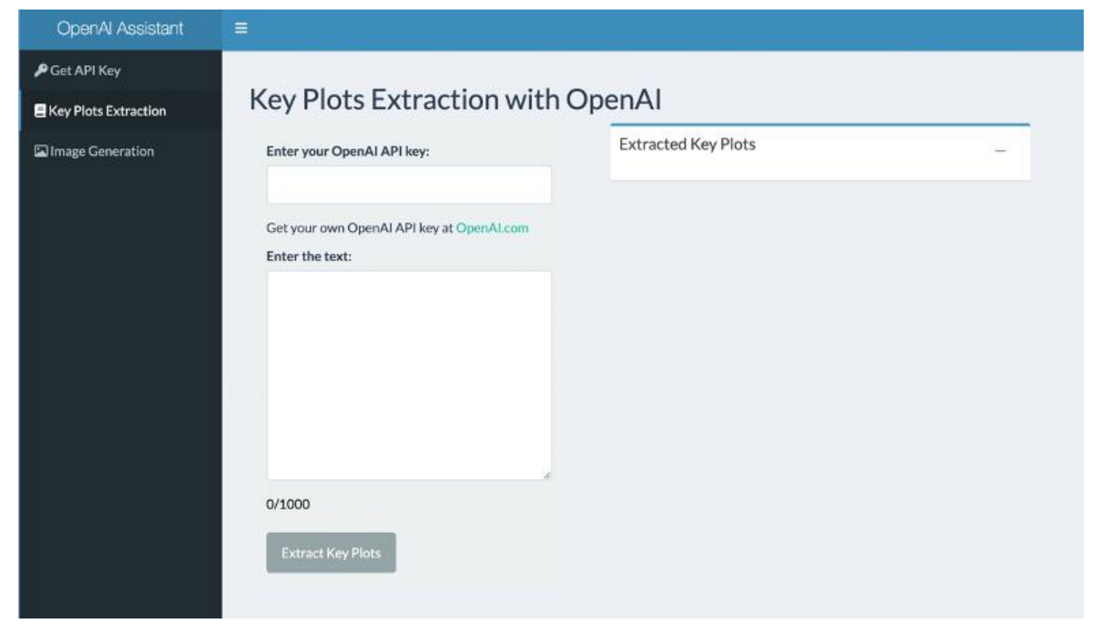
Get Your OpenAI API Key

Follow these steps to get your OpenAI API Key:

1. Visit the OpenAI platform website and sign up or log in.
2. If you are a new user, find 'Billing' on the left menu bar under 'Settings' and add the amount. Visit the OpenAI Pricing to know more about price charging.
3. If you are an old user or have already finished recharging, please find 'API keys' in the left menu and click 'Create a new secret key' to generate API keys.
4. Copy the API key when you use 'Key Plots Extraction' and 'Image Generation'.

Get API Key Page: On this page, users are given instructions to get their own API key for further use of Open AI.

Key Plots Extraction Page: This page is driven by GPT-4. On this page users should enter the API key they get from the previous step. Then, put the text they intend to convert to a image in the "*Enter the Text*" box. After clicking the "*Extract Key Plots*" button, we can get the extracted key plot based on the text we enter.



Key Plots Extraction with OpenAI

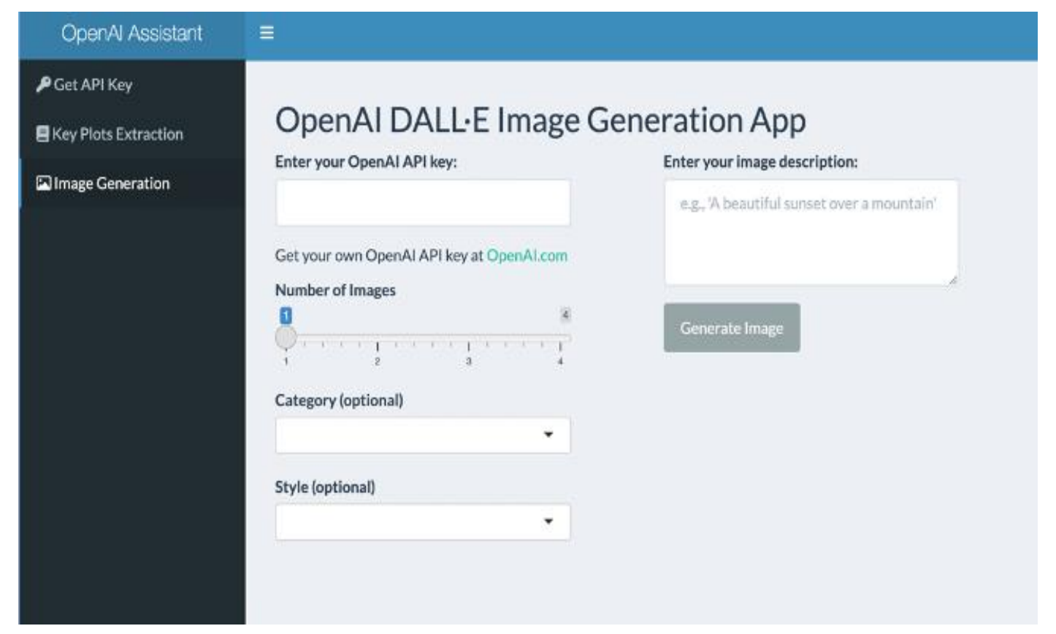
Enter your OpenAI API key:

Get your own OpenAI API key at OpenAI.com

Enter the text:

0/1000

Extracted Key Plots



OpenAI DALL·E Image Generation App

Enter your OpenAI API key:

Get your own OpenAI API key at OpenAI.com

Number of Images: 1

Category (optional)

Style (optional)

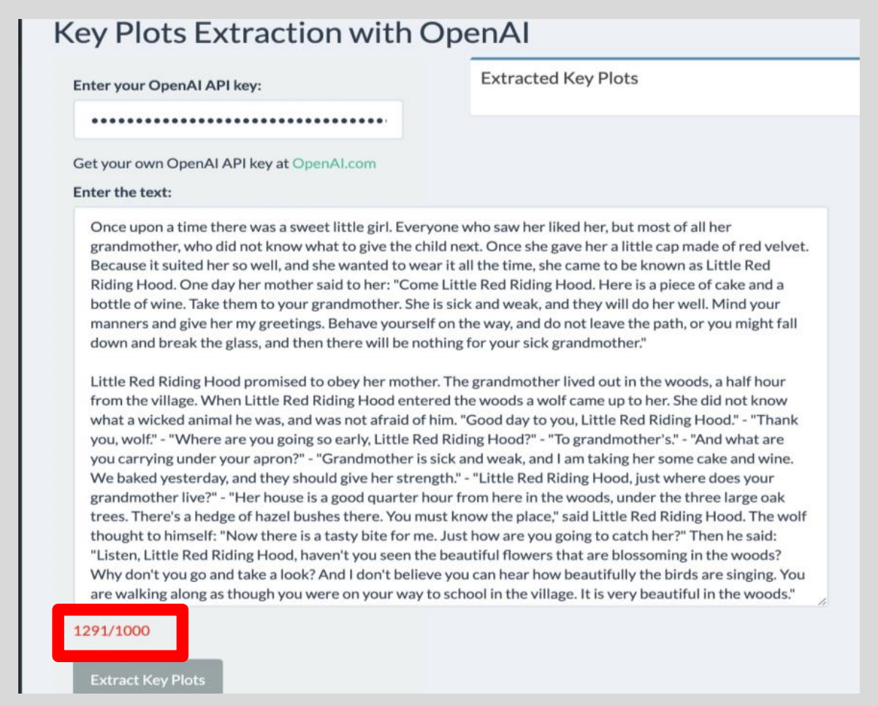
Enter your image description: e.g., 'A beautiful sunset over a mountain'

Image Generation Page: This page is driven by DALL·E 3. The user can copy and paste the first plot into the text box, select "*Line Drawing*" and "*Cartoon*" at the same time, and click "*Generate Picture*" to get the desired picture. If the picture is not satisfying, the user can generate it again and again by adding more descriptions in the text box and clicking the button. Every time the user click the button, AI will generate a new picture and show in the R shiny dashboard.

EVALUATION AND RESULT

To examine how accurate the application can convert textual stories into corresponding images. We design an experiment that is divided into four parts, aiming to find the most effective method to generate images that match the story content.

We take the classic fairy tale Little Red Riding Hood as an example, and the entire story is got from the website (https://americanliterature.com/childrens-stories/little-red-riding-hood). There are 1,442 words in the text, filled with dialog, description, and mental activity, including a number of longer paragraph.



Key Plots Extraction with OpenAI

Enter your OpenAI API key: ••••••••••••••••••••••••••••••••••

Get your own OpenAI API key at OpenAI.com

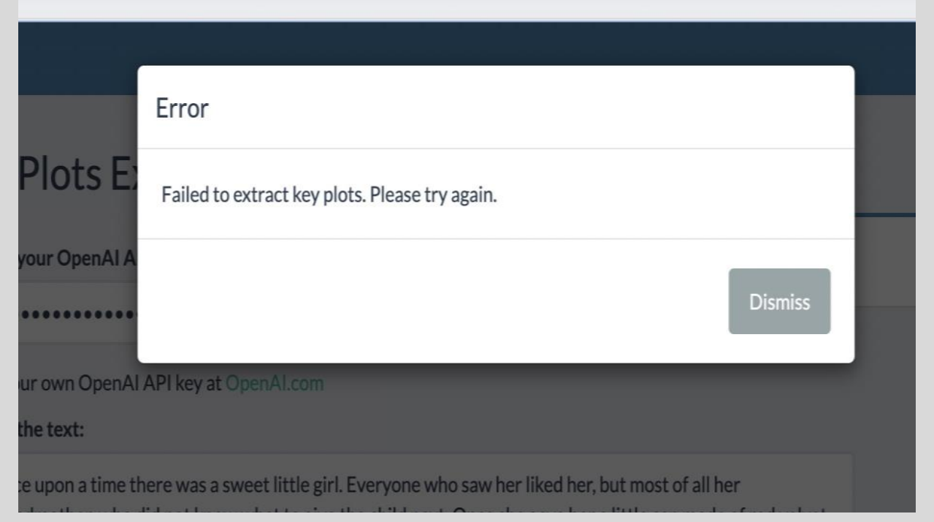
Enter the text:

Once upon a time there was a sweet little girl. Everyone who saw her liked her, but most of all her grandmother, who did not know what to give the child next. Once she gave her a little cap made of red velvet. Because it suited her so well, and she wanted to wear it all the time, she came to be known as Little Red Riding Hood. One day her mother said to her: "Come Little Red Riding Hood. Here is a piece of cake and a bottle of wine. Take them to your grandmother. She is sick and weak, and they will do her well. Mind your manners and give her my greetings. Behave yourself on the way, and do not leave the path, or you might fall down and break the glass, and then there will be nothing for your sick grandmother."

Little Red Riding Hood promised to obey her mother. The grandmother lived out in the woods, a half hour from the village. When Little Red Riding Hood entered the woods a wolf came up to her. She did not know what a wicked animal he was, and was not afraid of him. "Good day to you, Little Red Riding Hood." - "Thank you, wolf." - "Where are you going so early, Little Red Riding Hood?" - "To grandmother's." - "And what are you carrying under your apron?" - "Grandmother is sick and weak, and I am taking her some cake and wine. We baked yesterday, and they should give her strength." - "Little Red Riding Hood, just where does your grandmother live?" - "Her house is a good quarter hour from here in the woods, under the three large oak trees. There's a hedge of hazel bushes there. You must know the place," said Little Red Riding Hood. The wolf thought to himself: "Now there is a tasty bite for me. Just how are you going to catch her?" Then he said: "Listen, Little Red Riding Hood, haven't you seen the beautiful flowers that are blossoming in the woods? Why don't you go and take a look? And I don't believe you can hear how beautifully the birds are singing. You are walking along as though you were on your way to school in the village. It is very beautiful in the woods."

1291/1000

Extracted Key Plots

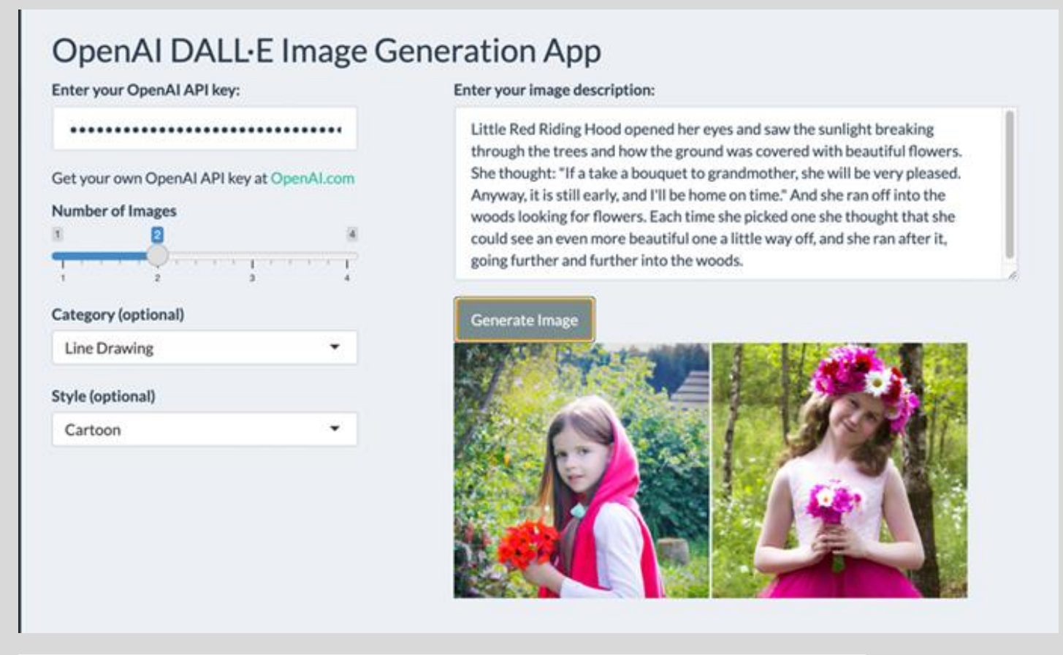


Error

Failed to extract key plots. Please try again.

The first part of the experiment is imputing full-length story text directly into DALL-E's dialogue box. The experiment failed due to DALL-E's input limitations, the entire story text exceeded the word limit, resulting in the inability to generate images.

The second part of the experiment is extracting keywords using "*Key Plot Extraction*", which extract keywords from ChatGPT to extract keywords, and then input the story keywords back to DALL-E. The experiment still failed. The extracted keywords might still be too numerous or not precise enough, causing DALL-E to be unable to process them effectively.



OpenAI DALL·E Image Generation App

Enter your OpenAI API key: ••••••••••••••••••••••••••••••••

Get your own OpenAI API key at OpenAI.com

Number of Images: 2

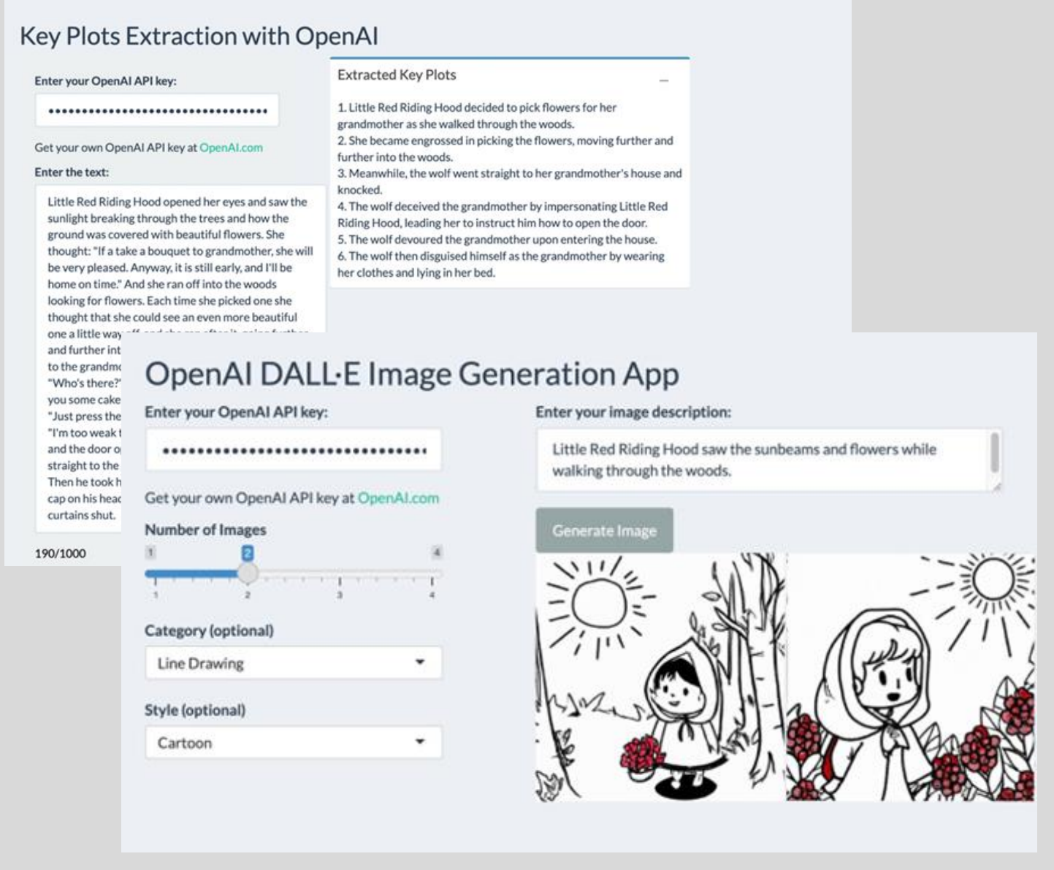
Category (optional): Line Drawing

Style (optional): Cartoon

Enter your image description:

Little Red Riding Hood opened her eyes and saw the sunlight breaking through the trees and how the ground was covered with beautiful flowers. She thought: "If a take a bouquet to grandmother, she will be very pleased. Anyway, it is still early, and I'll be home on time." And she ran off into the woods looking for flowers. Each time she picked one she thought that she could see an even more beautiful one a little way off, and she ran after it, going further and further into the woods.

In the third experiment the story was chopped into multiple segments, each of which was directly input into DALL-E. The experiment was not successful. Although DALL-E could handle shorter texts, the generated images were far from expectations and did not accurately reflect the story content (neither line drawings nor cartoons we selected from the side bar option).



Key Plots Extraction with OpenAI

Enter your OpenAI API key: ••••••••••••••••••••••••••••••••••

Get your own OpenAI API key at OpenAI.com

Enter the text:

Little Red Riding Hood opened her eyes and saw the sunlight breaking through the trees and how the ground was covered with beautiful flowers. She thought: "If a take a bouquet to grandmother, she will be very pleased. Anyway, it is still early, and I'll be home on time." And she ran off into the woods looking for flowers. Each time she picked one she thought that she could see an even more beautiful one a little way

190/1000

Extracted Key Plots

1. Little Red Riding Hood decided to pick flowers for her grandmother as she walked through the woods.
2. She became engrossed in picking the flowers, moving further and further into the woods.
3. Meanwhile, the wolf went straight to her grandmother's house and knocked.
4. The wolf deceived the grandmother by impersonating Little Red Riding Hood, leading her to instruct him how to open the door.
5. The wolf devoured the grandmother upon entering the house.
6. The wolf then disguised himself as the grandmother by wearing her clothes and lying in her bed.

OpenAI DALL·E Image Generation App

Enter your OpenAI API key: ••••••••••••••••••••••••••••••••

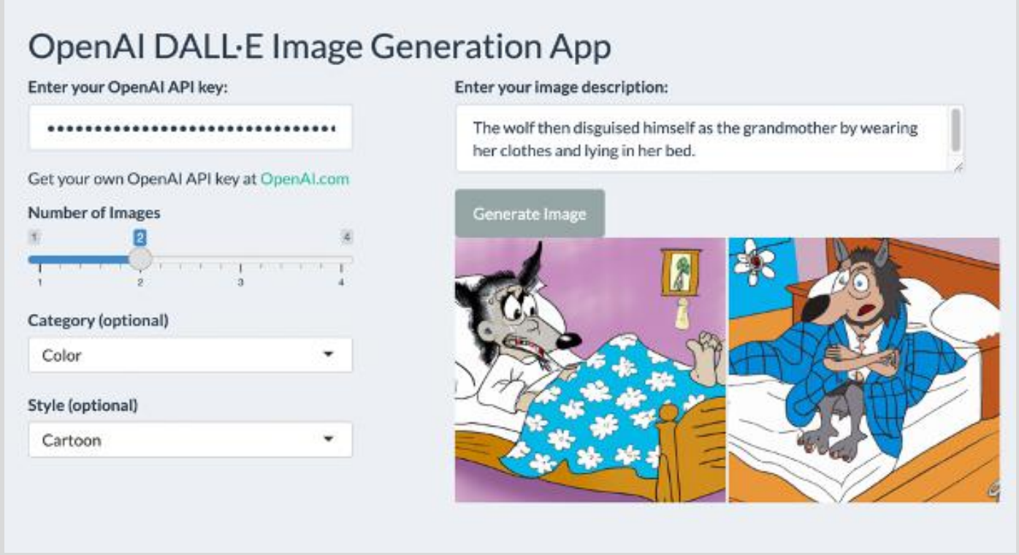
Get your own OpenAI API key at OpenAI.com

Number of Images: 2

Category (optional): Line Drawing

Style (optional): Cartoon

Enter your image description: Little Red Riding Hood saw the sunbeams and flowers while walking through the woods.



OpenAI DALL·E Image Generation App

Enter your OpenAI API key: ••••••••••••••••••••••••••••••••

Get your own OpenAI API key at OpenAI.com

Number of Images: 2

Category (optional): Color

Style (optional): Cartoon

Enter your image description: The wolf then disguised himself as the grandmother by wearing her clothes and lying in her bed.

In the fourth experiment, he story was chopped into segments, each processed through GPT to extract key plots, which were then input into DALL-E. The experiment was successful. After accurately extracting key plots with ChatGPT, DALL-E was able to generate images that matched the story content more closely based on these more refined and specific text prompts. Authors could choose any plot for further creation, providing clearer text prompts for AI.

CONCLUSIONS

In conclusion, our integration of R Shiny with OpenAI for our "Story Picture Generator" project underscores the vast potential at the crossroads of interactive platforms and AI models. The preliminary results show a transformative approach to visual storytelling. The AI's ability to interpret and translate textual content into images showed varying degrees of accuracy and effectiveness, largely dependent on the method of text processing and the complexity of the story content. Several key limitations and challenges were DALL-E's limitations in processing long text segments, and also inherent unpredictability in how AI understands and visualizes certain concepts, especially abstract or symbolic ones. Future experiments could delve into harnessing user preferences for image aesthetics, or introducing a feedback loop for continuous model optimization and so on. Moreover, embracing a multi-user interface might allow for collaborative story creation, paving the way for a new genre of collective digital narratives.

https://github.com/Yunlei02/Generate-Pictures-Tool