NewsNow: Automated News Retrieval System

A PROJECT REPORT

Submitted by

SEENUVASAN S (220701255)

in partial fulfillment for the course

OAI1903 - INTRODUCTION TO ROBOTIC PROCESS AUTOMATION

for the degree of

BACHELOR OF ENGINEERING

in

COMPUTER SCIENCE AND DESIGN

RAJALAKSHMI ENGINEERING COLLEGE

RAJALAKSHMI NAGAR

THANDALAM

CHENNAI – 602 105

NOVEMBER 2024

RAJALAKSHMI ENGINEERING COLLEGE
CHENNAI - 602105

BONAFIDE CERTIFICATE

Certified that this project report **"NewsNow: Automated News Retrieval System"** is the bonafide work of **"SEENUVASAN S (220701255)"** who carried out the project work for the subject OAI1903 - Introduction to Robotic Process Automation under my supervision.

SIGNATURE   SIGNATURE

DR. P. KUMAR, M.E., Ph.D.,   **Dr. N. Durai Murugan, M.E., Ph.D.,**
HEAD OF THE DEPARTMENT,   **SUPERVISOR,**
Professor and Head,   Assistant Professor (SG),
Department of   Department of
Computer Science and Engineering,   Computer Science and Engineering,
Rajalakshmi Engineering College,   Rajalakshmi Engineering College,
Rajalakshmi Nagar,   Rajalakshmi Nagar,
Thandalam,   Thandalam,
Chennai - 602105.   Chennai – 602105.

Submitted to Project and Viva Voce Examination for the subject OAI1903 - Introduction to Robotic Process Automation held on _____.

Internal Examiner   **External Examiner**

ABSTRACT

This project, **"NewsNow: Automated News Retrieval System,"** leverages Robotic Process Automation (RPA) to simplify and streamline the process of retrieving the latest news headlines. Designed for efficiency and user-friendliness, the system begins by taking user input for a specific topic to search for, as well as their email address, through an input dialog box. Using this information, the bot performs an automated Google search for recent news articles related to the topic.

The bot then filters the search results to identify articles published within the last few hours, ensuring only the most recent updates are considered. Once a suitable article is found, the bot extracts its headline and URL. Finally, this information is sent to the user via email, providing them with timely news on their chosen topic without requiring manual effort.

This project demonstrates the practical application of RPA in automating routine tasks, saving time, and enhancing productivity. It is particularly useful for individuals and professionals who need real-time updates on specific topics without the hassle of manually searching the web. The system showcases the potential of RPA in bridging automation and information accessibility efficiently.

ACKNOWLEDGEMENT

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF ABBREVIATIONS

ABBREVIATION	ACCRONYM
RPA	Robotic Process Automation
AI	Artificial Intelligence
API	Application Programming Interface
CV	Computer Vision
OCR	Optical Character Recognition

CHAPTER 1

INTRODUCTION

1.1 INTRODUCTION

In the digital age, staying informed about the latest news is more important than ever. With an overwhelming amount of information available online, it can be challenging to sift through countless articles and news sources to find relevant updates. Traditional methods of manually searching for news articles can be both time-consuming and inefficient, especially when there is a need for real-time information. This is where Robotic Process Automation (RPA) comes into play.

RPA is a technology that uses software robots or "bots" to automate repetitive tasks, which would otherwise require human intervention. These bots are designed to interact with digital systems and execute processes in a way that mimics human actions, but faster and more accurately. The **"NewsNow: Automated News Retrieval System"** leverages the power of RPA to streamline the process of news retrieval, making it more efficient and user-friendly.

The system is designed to take input from the user, asking them to specify a topic they are interested in and their email address. Based on this input, the bot conducts a Google search for recent news articles on the given topic. The bot then applies a filter to ensure that only articles published within a specific timeframe (the last few hours) are considered. This filter is based on the presence of the phrase "hours ago" in the article's timestamp.

Once the latest news headline is identified, along with the article's URL, the bot automatically sends this information to the user's email address. This process eliminates the need for the user to manually search for news,

keeping them up to date with the most relevant information in a fraction of the time. The system is ideal for professionals, researchers, and anyone who needs quick and accurate news updates without having to dedicate time to manual searches. The **"NewsNow"** system not only enhances efficiency but also demonstrates the significant potential of RPA in automating real-world tasks.

1.2 OBJECTIVE

The objective of the **"NewsNow: Automated News Retrieval System"** is to automate the process of retrieving and delivering the latest news updates to users based on their specific interests. By leveraging Robotic Process Automation (RPA), this system eliminates the need for users to manually search for news articles, saving both time and effort. The system is designed to perform an automated Google search for news articles on a specified topic, filter out outdated content, and deliver the most recent headlines along with their URLs to the user via email. This automation ensures that users receive timely, relevant news with minimal intervention. The key objectives of the system are:

1. To allow users to input a topic of interest and their email address.
2. To automate the process of searching for the most recent news articles related to the topic.
3. To filter news articles to include only those published within the last few hours.
4. To send the most recent headline and article link to the user's email address.
5. To provide an easy-to-use interface for users to receive automated updates.

1.3 EXISTING SYSTEM

Currently, there is no automated system in place for retrieving and sending timely news updates based on a user's specified topic. Users must rely on manual methods such as searching news websites, using news aggregator apps, or subscribing to email newsletters. These methods, while useful, can be time-consuming and often fail to provide the latest, most relevant updates in real-time. Additionally, users may struggle with filtering out outdated content and may miss crucial updates during busy hours.

1.4 PROPOSED SYSTEM

The proposed **"NewsNow"** system leverages RPA to automate news retrieval and delivery. Upon receiving input from the user through a simple dialog box, the bot performs an automated Google search for the most recent news articles on the specified topic. It then filters articles to ensure that only those published within a specified timeframe (e.g., the last few hours) are included. The system then extracts the latest headline and URL of the relevant article and sends it directly to the user's email. This eliminates the need for manual searches and ensures that users receive real-time, curated news updates without effort. This system improves efficiency and ensures timely, relevant information is delivered with minimal user involvement.

CHAPTER 2

LITERATURE REVIEW

2.1 Survey on Robotic Process Automation (RPA) in News Retrieval and Automation:

Robotic Process Automation (RPA) has found applications across various domains, with its potential to automate repetitive tasks and enhance efficiency. In the context of news retrieval, RPA is used to streamline processes like data collection, filtering, and notification generation, ensuring that users stay updated with minimal effort. RPA technology is especially beneficial in tasks like searching for specific information on the internet, sorting data based on relevance, and automating the process of delivering information to end-users. Below are some key studies and applications of RPA in automated information retrieval:

[1] A study on the use of RPA in content aggregation highlights how businesses use RPA bots to automate the process of collecting relevant news and updates from various online sources. The study emphasizes how RPA technology helps in gathering specific types of information in real-time, such as news headlines and URLs, providing users with timely updates without manual intervention.

[2] A research paper on RPA in the information services industry discusses the automation of various processes, such as monitoring news and sending alerts based on predefined criteria. It shows how RPA can automate the process of filtering and identifying articles based on recency and relevance, ensuring that users receive only the most current and relevant news updates.

2.2 Survey on Automated News Delivery Systems:

Automated news delivery systems have become increasingly popular as they allow users to receive updates without actively searching for them. These systems utilize various automation technologies, including RPA, to collect, filter, and deliver news directly to users. The challenge lies in ensuring that the news delivered is not only relevant but also timely. Several studies explore the challenges and benefits of automated news delivery systems:

[3] A research paper explores how automated systems are employed to fetch and deliver news updates in real-time. The paper discusses various filtering techniques used to sort articles based on their publication time and relevance. It highlights the importance of delivering the most recent news to users and ensuring that content is free from outdated information.

[4] A study focuses on the effectiveness of RPA in automating news searches on the internet. The paper explains how RPA bots are used to search for specific keywords, filter results based on recency (e.g., articles published within hours), and deliver the relevant headlines and links to users via email. It emphasizes the importance of precision in filtering and ensuring that outdated articles are excluded.

2.3 Survey on Email Notification Systems for Automated Information Delivery:

Email notifications are one of the most effective ways to deliver automated updates to users, ensuring that they receive important information in a timely manner. Automated email systems have been used extensively to send alerts and notifications based on user preferences. In the context of news retrieval, sending real-time updates via email is an essential part of the system. Below are studies related to automated email notification systems:

[5] A study examines the use of automated email systems in news delivery. It focuses on how automated systems can send personalized news alerts based on the topics of interest specified by users. The research highlights how email notifications can be tailored to deliver concise, timely news directly to users, improving engagement and efficiency.

[6] A paper on the use of email notification systems in automated content delivery discusses the technical challenges of integrating RPA with email systems. It highlights the importance of ensuring that notifications are sent quickly and accurately, especially when dealing with real-time data, such as breaking news updates.

2.4 Summary of the Intersection of RPA, News Retrieval, and Email Notification Systems:

The proposed **"NewsNow: Automated News Retrieval System"** leverages RPA to automate the process of searching for relevant news articles based on user input. The integration of RPA with automated email systems ensures that users receive timely and relevant news directly to their inbox. This project builds on existing research and innovations in the field of automation, enhancing the ability to deliver real-time news updates with minimal user effort.

The project merges RPA with email notifications, allowing users to specify a topic and receive the latest news headlines and links through an automated system. This integration is a significant step forward in providing a practical solution for users who need up-to-date information without engaging in manual searches. By automating the process of retrieving and delivering news, the project contributes to the ongoing advancements in RPA-based systems for automated information delivery.

CHAPTER 3

SYSTEM DESIGN

3.1 SYSTEM FLOW DIAGRAM

A flowchart is a type of diagram that represents an algorithm, workflow or process. The flowchart shows the steps as boxes of various kinds, and their order by connecting the boxes with arrows. This diagrammatic representation illustrates a solution model to a given problem. The system flow diagram for this project is in Fig. 3.1.
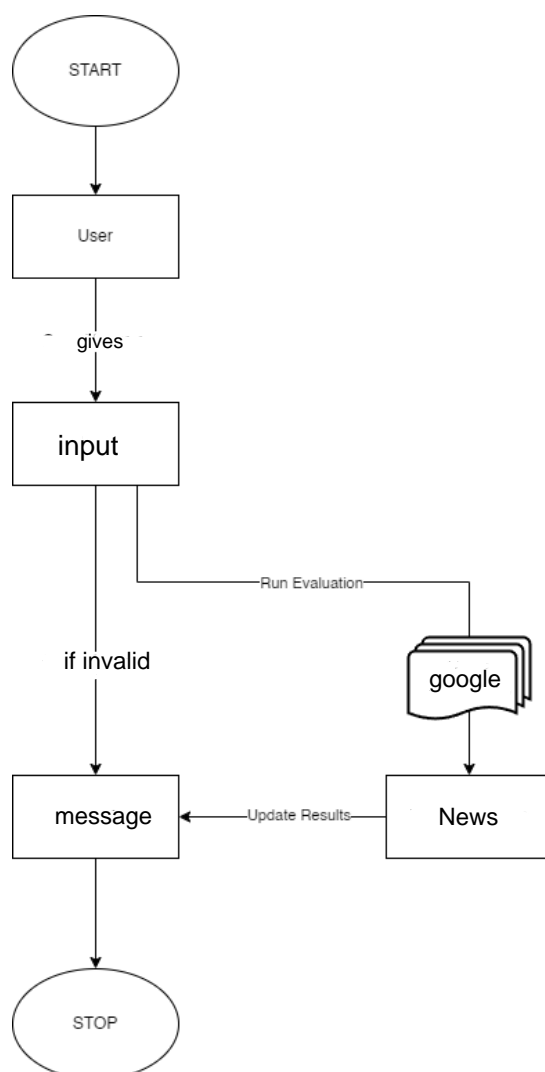

Fig 3.1 System Flow Diagram

3.2 ARCHITECTURE DIAGRAM

An architecture diagram is a graphical representation of a set of concepts, that are part of an architecture, including their principles, elements and components. The architecture diagram for this project is in Fig. 3.2.
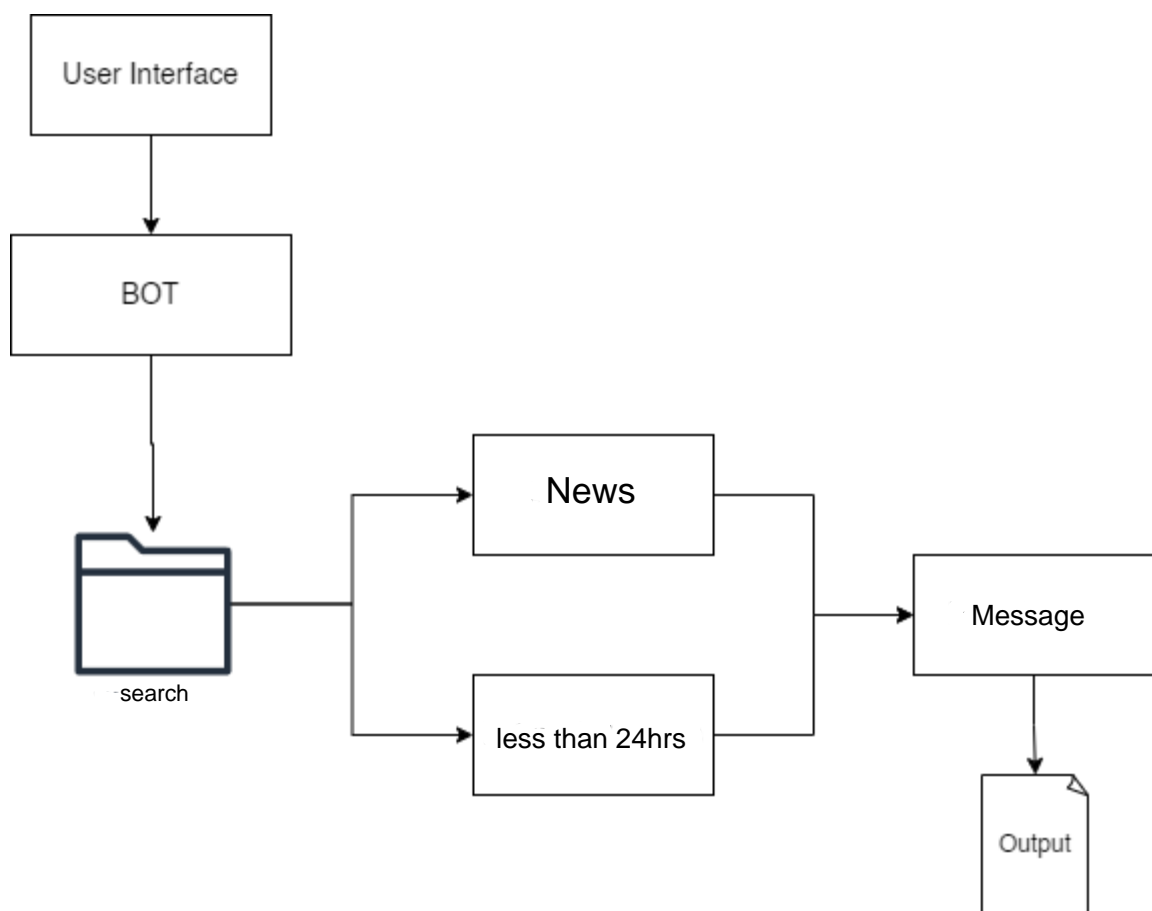

Fig 3.2 Architecture Diagram

3.3 SEQUENCE DIAGRAM

A sequence diagram is a type of interaction diagram because it describe and s how in what order a group of objects works together. The sequence diagram for this project is in Fig. 3.3.
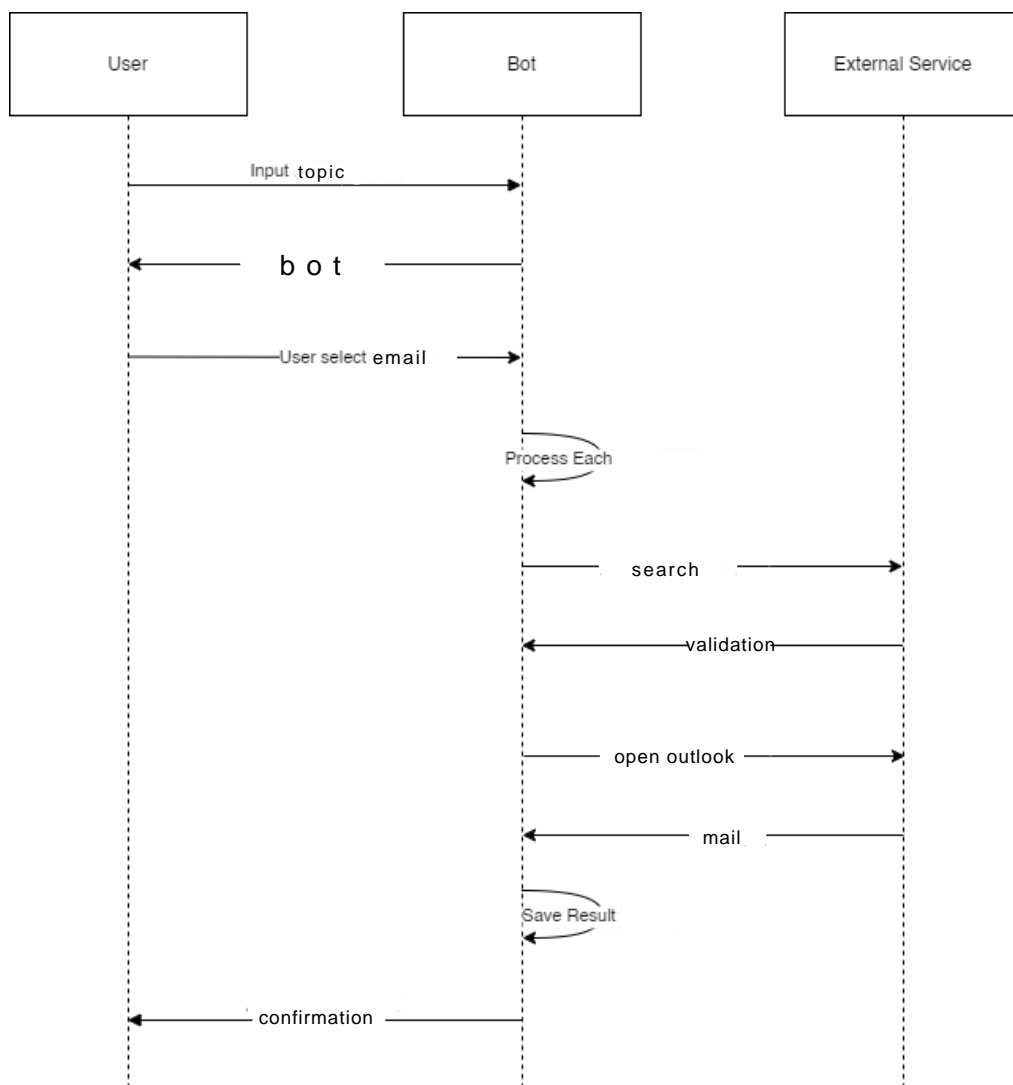

Fig 3.3 Sequence Diagram

11

CHAPTER 4

PROJECT DESCRIPTION

The **"Smart News Retrieval and Delivery Bot"** is an advanced Robotic Process Automation (RPA) project aimed at automating the process of news retrieval and delivery based on user-specified topics. Developed using UiPath, this intelligent bot is designed to address the need for timely and relevant news updates by efficiently searching for the most recent news articles on a given subject, filtering outdated content, and sending the latest headlines along with their URLs directly to the user's email.

4.1. MODULES:

4.1.1. INPUT HANDLING AND INITIALIZATION:

4.1.1.1. Get Topic:

- Receive user input for the topic.

4.1.1.2. Get Email:

- Allow the user to give email id to which the recent news will be sent for the topic.

4.1.1.3 Confirmation:

- Dynamically prints a Message named "News Sent".

4.1.2 CONTENT ANALYSIS:

4.1.2.1 Topic Search and News Retrieval:

- **Iterate through search results**: The bot begins by accepting the user's input, which includes the topic to search and the user's email. After the topic is received, it performs a search using Google to gather the most relevant news articles for that topic.

4.1.2.2 Email Delivery:

- **Send to the user's inbox**: The email is then sent directly to the user, ensuring that they receive timely and relevant news updates on their specified topic without having to manually search for it.

4.1.3 RESULT MANAGEMENT:

4.1.3.1 Result Storage:

- Systematically update the data table with the results.

4.1.3.2 Real-time Update:

- Display real-time updates of the integrity verification process.

4.1.4 COMPLETION AND REPORTING:

4.1.4.1 Completion Message:

- Conclude the operation with a message indicating the successful completion of the integrity verification task.

CHAPTER 5

OUTPUT SCREENSHOTS

Fig 5.1 – Input Dialog

Fig 5.2

Topic : news Inbox ×

SEENUVASAN S <seenuseenu28156401@outlook.com>
to me ▾

Latest on news-> Exit Poll Results 2024 Highlights: BJP+ Ahead In Maharashtra, Tight Race In Jharkhand, Predict Exit Polls
URL-> https://www.ndtv.com/india-news/exit-poll-results-2024-maharashtra-and-jharkhand-live-updates-7063591
Thank you!

Best Regards,
NewsNow Bot.

← Reply   → Forward   ☺
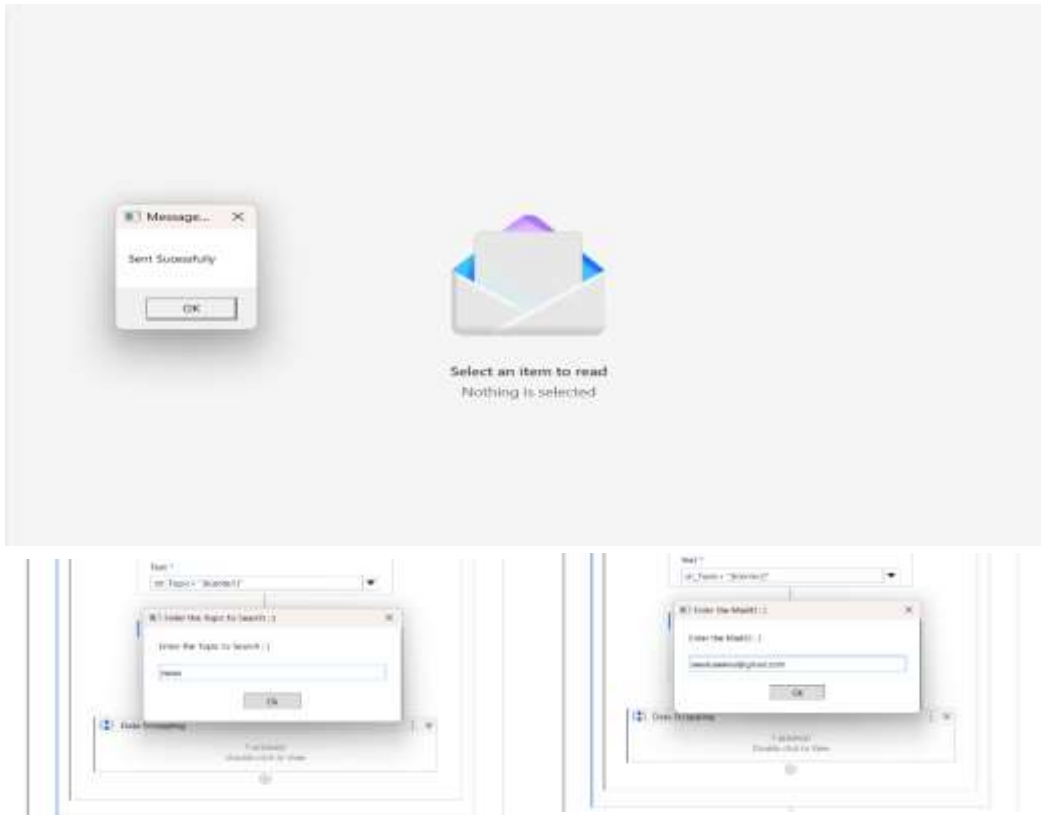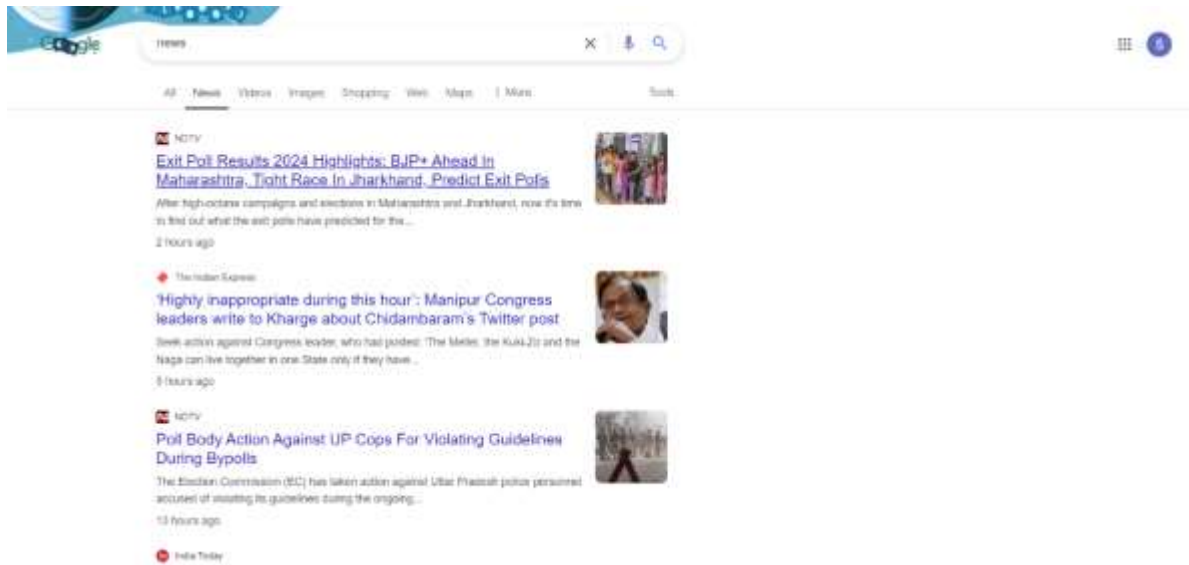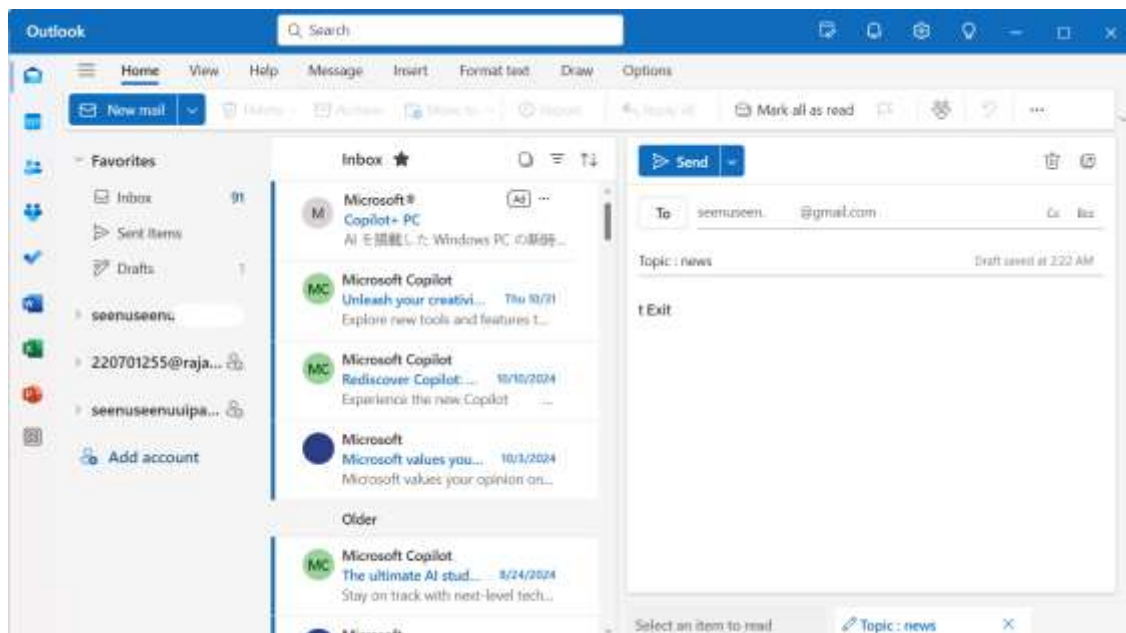
Fig 5.3

CHAPTER 6

CONCLUSION

The **"Smart News Retrieval and Delivery Bot"** represents an innovative use of Robotic Process Automation (RPA) technology to automate the process of news retrieval and delivery based on user-defined topics. With the rapid pace of information dissemination, staying updated on relevant topics is increasingly challenging. Traditional methods of searching for news can be time-consuming, and manually sifting through outdated or irrelevant articles can be inefficient. This project aims to solve these problems by providing a seamless and automated solution to keep users informed in real-time.

The bot efficiently takes user input—such as the topic of interest and the recipient's email address—through a simple dialog box. It then performs an automated search on Google, retrieves the most recent news articles, filters them based on the "hours ago" condition, and ensures that the content delivered is up-to-date and relevant. By leveraging RPA, this process is streamlined and automated, significantly reducing the time and effort needed to stay informed on various topics.

One of the key benefits of this project is the automation of news retrieval and email delivery. This allows users to bypass the need for manual searches, enabling them to receive the most current news directly in their inbox. The efficiency and time-saving nature of the bot are especially beneficial for users who need timely information but lack the time to sift through multiple sources manually. Furthermore, the integration of automated email notifications provides a hands-off experience for the user, as they receive news updates without needing to open multiple websites or

apps.

The project's effectiveness in delivering relevant and timely news is rooted in its intelligent filtering mechanism. By only including articles published in the past few hours, the bot ensures that users are not burdened with outdated or irrelevant content. This feature is particularly important in a fast-moving digital world, where news can quickly become obsolete. The result is an efficient, user-friendly solution that offers the most recent headlines, making it a valuable tool for individuals and organizations who need to stay updated on specific topics.
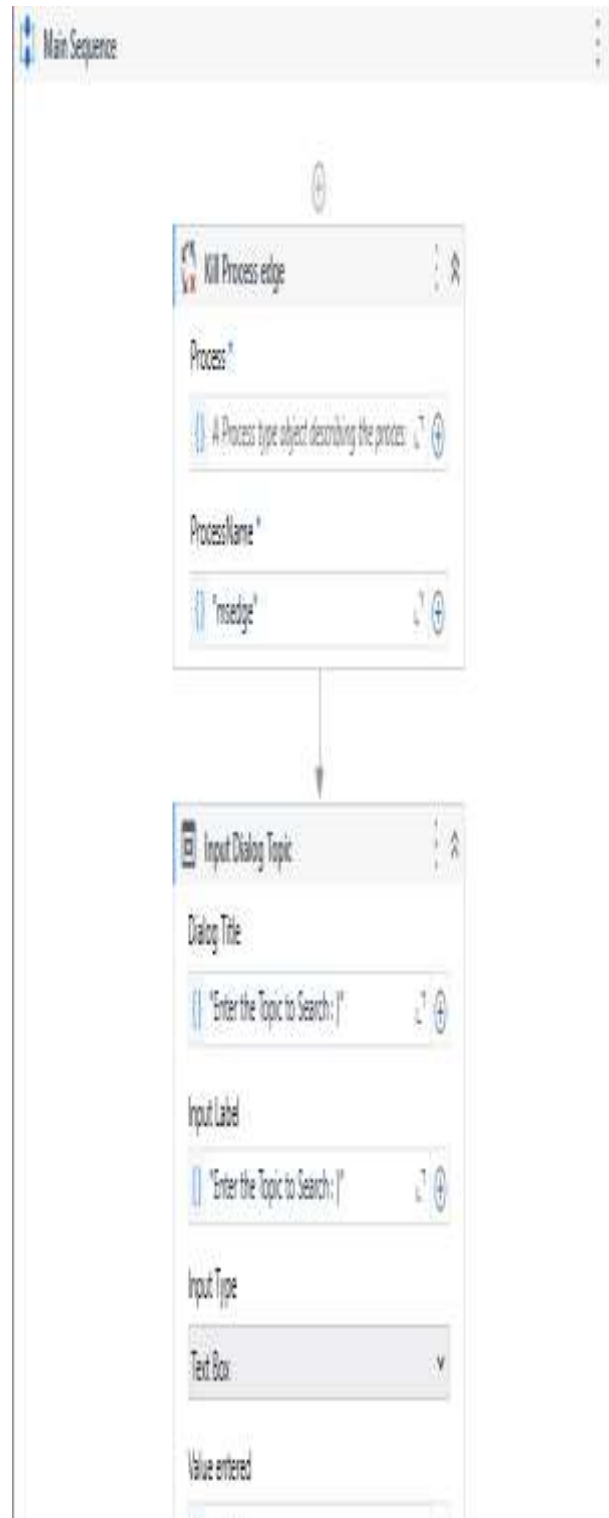
The development of this project highlights the power and potential of Robotic Process Automation in simplifying and automating routine tasks. RPA has the ability to improve efficiency, reduce human intervention, and enhance the user experience, all of which are crucial for today's information-driven society. By automating the process of news retrieval and email delivery, the project not only saves users valuable time but also ensures that they receive high-quality, timely information.
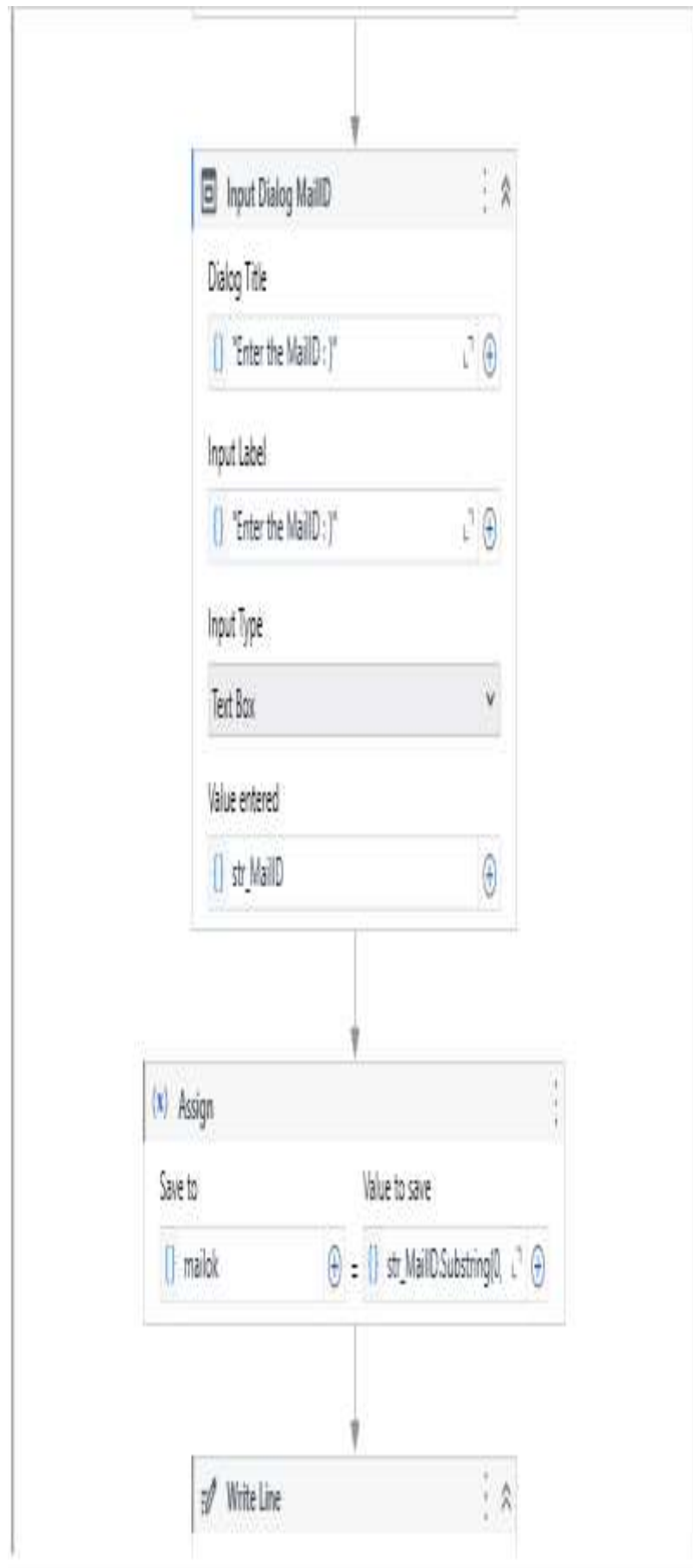
Looking ahead, there are several potential enhancements that could further improve the functionality of the **"Smart News Retrieval and Delivery Bot."** One potential feature is the ability to retrieve news from multiple sources, expanding the range of news articles available to the user. Additionally, incorporating machine learning algorithms could improve the bot's ability to discern between the relevance of news articles, providing users with even more personalized and tailored content. Another possible enhancement could include integrating additional filtering options based on the user's preferences, such as news sources, keywords, or article length.

Furthermore, expanding the bot's capabilities to handle multiple topics at once or automating news retrieval for a set of topics over a specified time frame could benefit users who want to stay updated on a range of issues or industries. The project could also explore the option of integrating with other platforms, such as messaging apps, to send the news in a more instant.

APPENDIX

PROCESS WORK FLOW

Input Dialog MailID

Dialog Title

"Enter the MailID :)"

Input Label

"Enter the MailID :)"

Input Type

Text Box

Value entered

str_MailID

Assign

Save to		Value to save
mailck	=	str_MailID.Substring(0,

Write Line
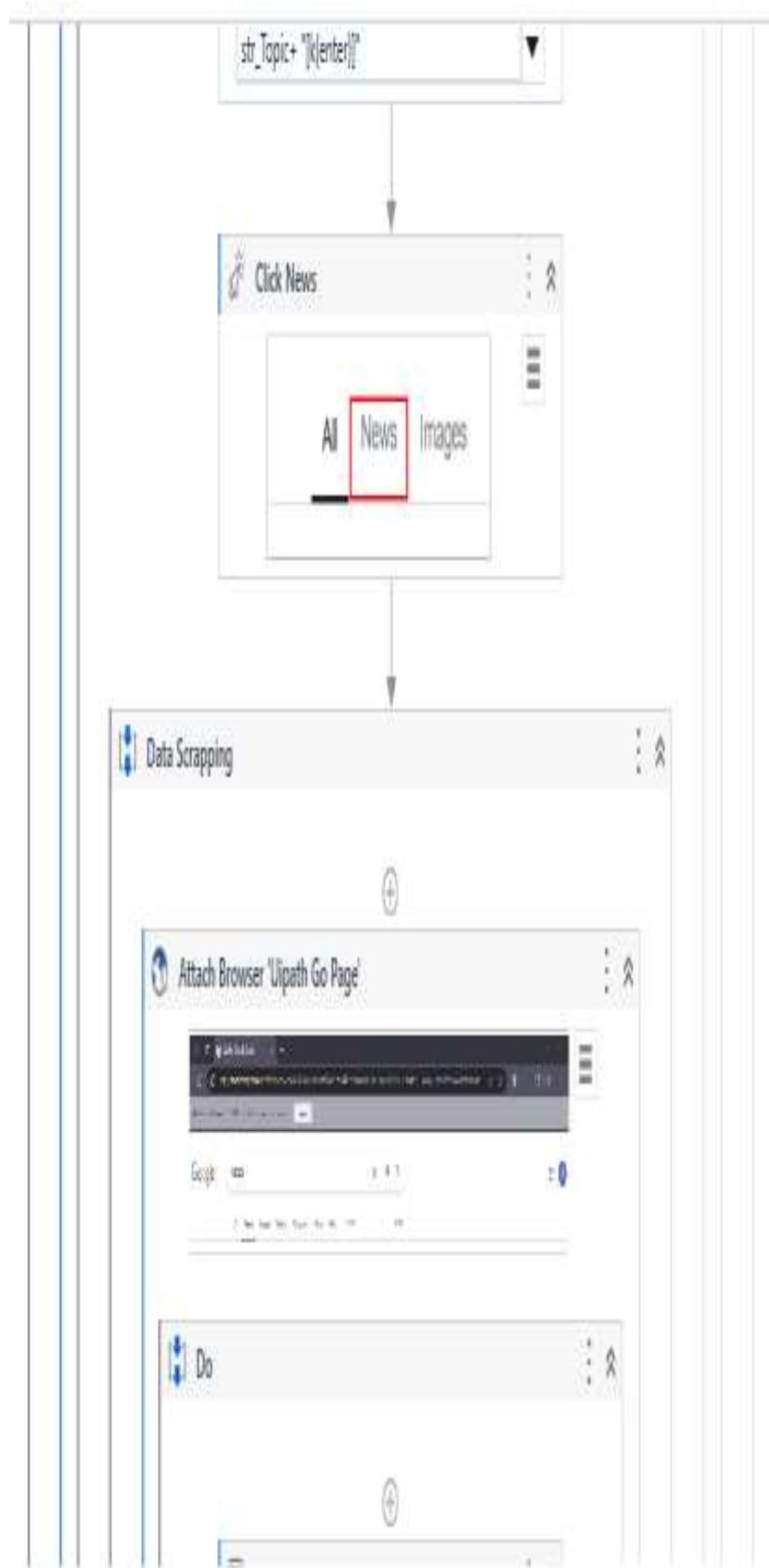
Write Line

Text

{ } str_MailID

Open Browser

Url *

"www.google.com"

Do

Type Into Search Box

Text *

str_Topic+ "[k(enter)]"

20

str_Topic+ "[k(enter)]" ▼

Click News

All News Images

Data Scrapping

⊕

Attach Browser 'Uipath Go Page'

Google

Do

⊕

Write Line

Text

{} str_Time_First_News

If

Condition *

{} str_Time_First_News.ToLower.Contains("ho

Then

Open Application 'Chrome_RenderWidgetHost...'

Outlook (new)
App

Do

Open Application 'Chrome_RenderWidgetHost...'

Outlook (new)
App

Do

Delay
Duration *

00h 00m 05.000s

Click 'grouping'

New mail

Type Into 'editable text To'

TI Type Into 'editable text To'

To

Text *

str_MailID ▼

TI Type Into 'editable text'

...

Add a subject

Text *

"Topic : "+str_Topic ▼

TI Type Into 'editable text'

TI Type Into 'editable text'

Text *

"Latest on "+str_Topic+"-> "+ExtractDataTable.Row ▼

Click 'push button Send'

▷ Send ⌄ From:

⏱ Delay

Duration *

00h 00m 03.000s ⌃⌄ ⊕

Duration *

00h 00m 03.000s

Message Box

Text *

{} "Sent Sucessfully"

↳ Else

Message Box

Text *

{} "error occured"
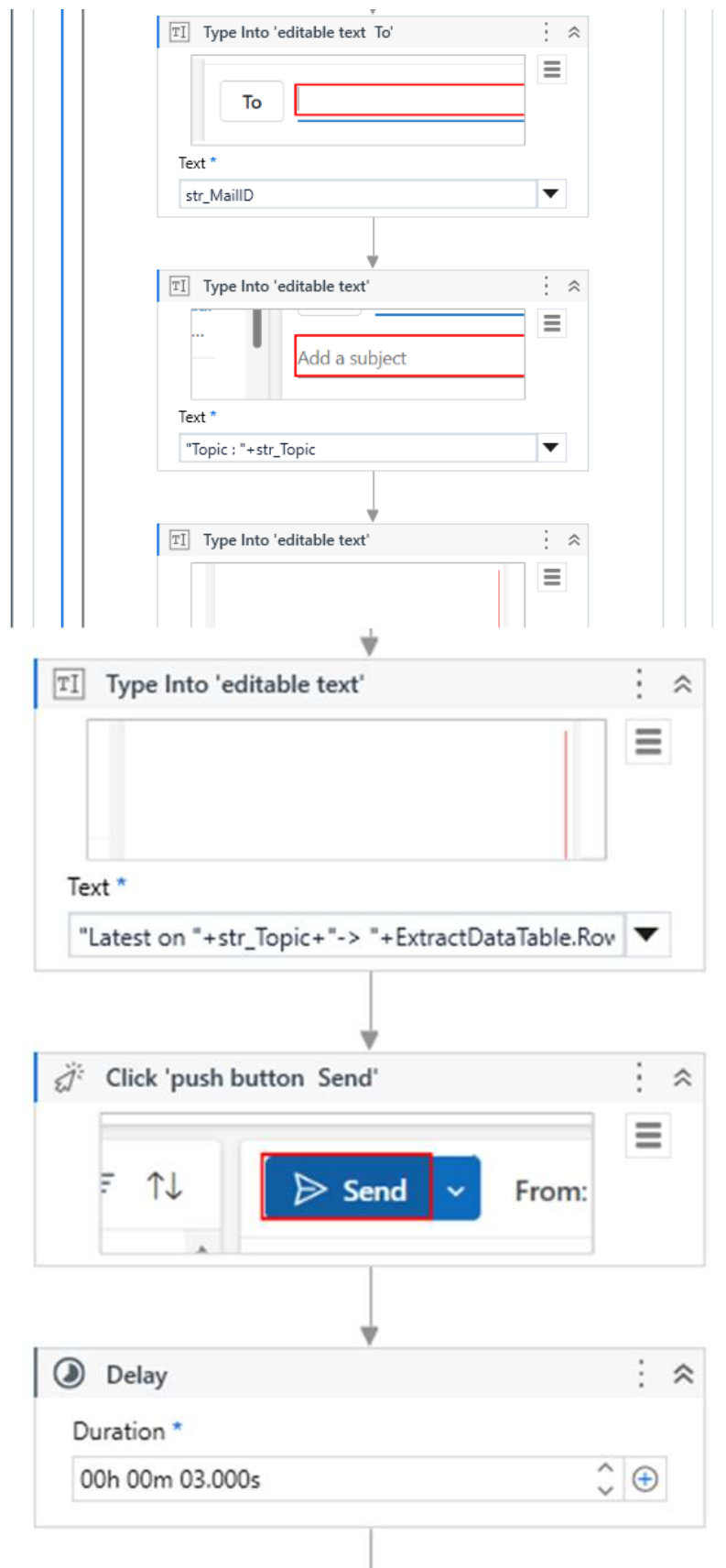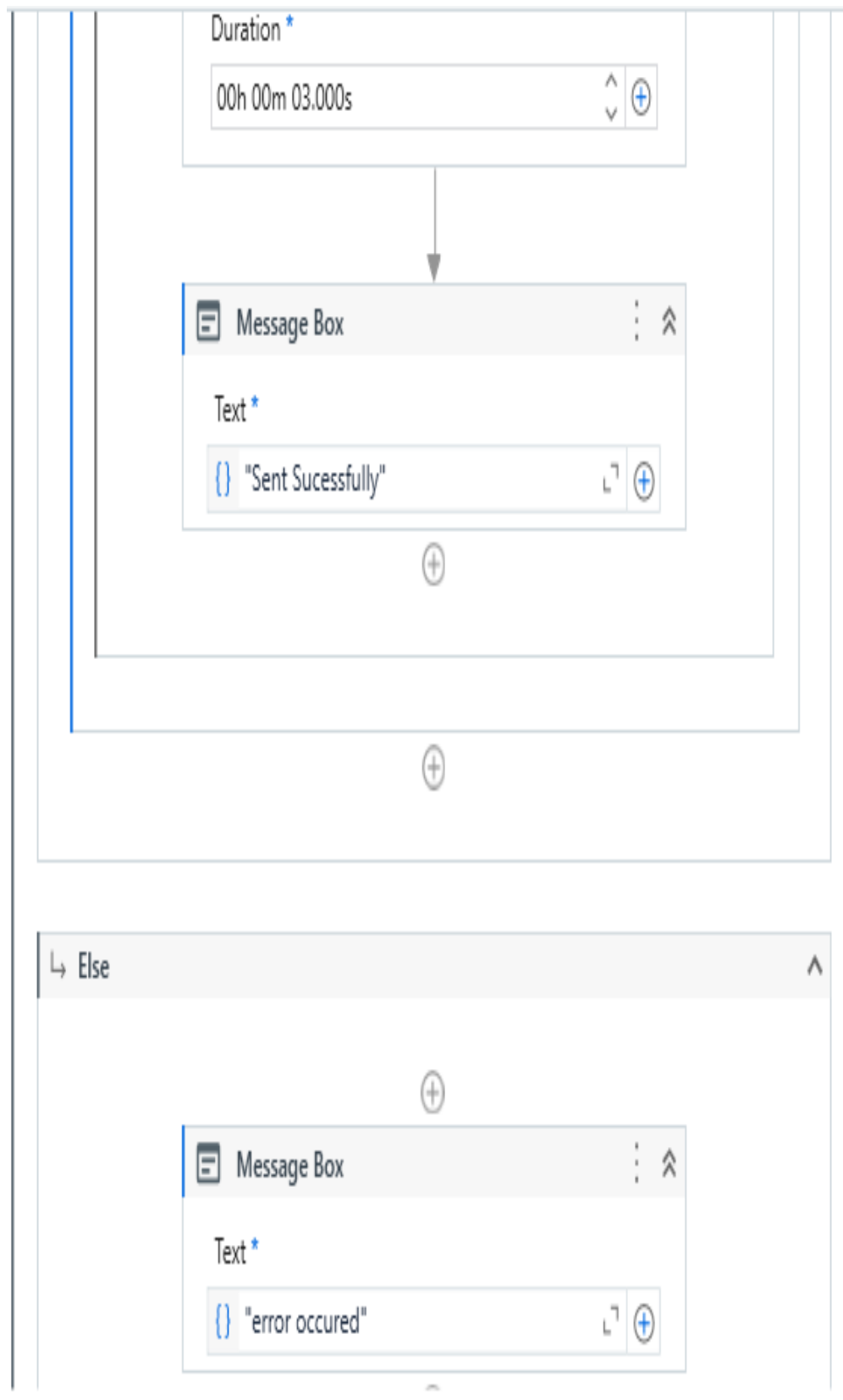
REFERENCES

[1] Kuppusamy, Palanivel & Joseph K, Suresh. (2020). Robotic Process Automation to Smart Education. 3775.

[2] Patil, Dr & Mane, Vinod & Patil, Dr. (2019). Social Innovation in Education System by using Robotic Process Automation (Rpa). International Journal of Innovative Technology and Exploring Engineering. 8. 3757-3760. 10.35940/ijitee.K2148.0981119.

[3] Elkhatat, A.M., Elsaid, K. & Almeer, S. Evaluating the efficacy of AI content detection tools in differentiating between human and AI-generated text. *Int J Educ Integr* **19**, 17 (2023). https://doi.org/10.1007/s40979-023-00140-5

[4] H. Alamleh, A. A. S. AlQahtani and A. ElSaid, "Distinguishing Human-Written and ChatGPT-Generated Text Using Machine Learning," 2023 Systems and Information Engineering Design Symposium (SIEDS), Charlottesville, VA, USA, 2023, pp. 154-158, doi: 10.1109/SIEDS58326.2023.10137767.

[5] Tomáš Foltýnek, Norman Meuschke, and Bela Gipp. 2019. Academic Plagiarism Detection: A Systematic Literature Review. ACM Comput. Surv. 52, 6, Article 112 (November 2020), 42 pages. https://doi.org/10.1145/3345317

[6] H. A. Chowdhury, D. K. Bhattacharyya, "Plagiarism: Taxonomy, Tools and Detection Techniques", 19th National Convention on Knowledge, Library and Information Networking, 2018.