

Robotic Process Automation Driven - Email Automation and Spam Detection

Dr.N.Mohanapriya

Associate Professor/CSE,
Department of Computer Science and Engineering
Vivekanandha College of Engineering for Women
Namakkal, India

K.Sivapriya

Assistant Professor,
Department of Computer Science and Engineering
Vivekanandha College of Engineering for Women Namakkal,
India

V.Padmapriya, B.Shibidharshana, K.Rajnandini

UG Scholar

Department of Computer Science and Engineering
Vivekanandha College of Engineering for Women
Namakkal, India

Abstract: In order to improve email processes, this study explores the use of robotic process automation (RPA), with an emphasis on effective spam identification and automation. In a time when email correspondence is essential to organizational processes, integrating RPA appears to be a way to boost output and cut down on human interaction. Using robotic process automation, rule-based email management chores can be automated. Routine tasks like email sorting, categorization, and response are accelerated by this automation. Organizations can enhance their operational efficiency and accuracy in email-related duties by reducing the dependence on manual labour for repeated procedures. Additionally, the study tackles the problem of spam identification, highlighting the function of RPA in protecting email systems from unsolicited messages. Dynamic spam filters that adjust to new spam strategies are developed in part via RPA. As a result, the spam detection system becomes more effective and responsive, guaranteeing a neater and better-organized email environment. This article concludes by emphasizing the need of utilizing robotic process automation for spam detection and email automation. It also provides insights into how businesses can improve the overall effectiveness of email management procedures and optimize their email workflows.

Keywords—E-Mail Automation, Dynamic spam filters, Robotic Process Automation (RPA), Categorization, Responsive system, UiPath.

I. INTRODUCTION

As of 2019, 3.9 billion people globally used email apps, with a projected increase to 4 billion by 2020. Managing the substantial daily volume of emails, even in midsize businesses, demands significant storage [1]. Addressing this challenge involves an escalating demand for automation, with Robotic Process Automation (RPA) emerging as a practical solution. RPA efficiently handles email processing, allowing human resources to focus on more critical tasks [1]

New digital tools like Robotic Process Automation (RPA) improve organizational processes, cutting costs and enhancing efficiency. RPA, seen as lightweight IT innovation, automates manual tasks using trained robots [2]. The study examines successful RPA adoption in ten case studies, revealing key success factors. Despite potential, initial RPA programs face a 30-50% failure rate, prompting a thorough analysis in the article [3]. Robotic Process Automation (RPA) utilizes software robots to automate repetitive tasks, emulating human behavior without infrastructure changes. It enhances efficiency, saves money, and time, particularly in rule-based operations with structured data. RPA's iterative learning capability extends to discovering efficient approaches for scalable tasks, driving interest in broader applications for management scenarios [4].

To optimize workflows and enhance accuracy, suggested technologies include Optical Character Recognition (OCR), computer vision, and Robotic Process Automation (RPA). RPA 's affordability has spurred its adoption in transportation logistics, highlighting the demand for improved solutions [5].

Email usage is now more common than half of all people on the planet; it is projected to reach 3.9 billion users by 2019 and 4 billion by 2020. Based on a 3 percent annual growth rate, the anticipated total email users worldwide in 2023 will be close to 4.3 billion. Even with improvements in big data storage, handling the large amount of emails is still difficult, especially when you take into account the variety of information that may call for specialized knowledge in different fields. In order to deal with this, there's an increasing interest in email processing automation, which makes it possible to prioritize, segregate, and handle responses efficiently using tools like robotic process automation (RPA) [6]. Email remains a common form of contact in both personal and professional contexts, providing the thrill of immediate answers. Instantaneous responses are desired for efficiency and enjoyment in our tech-driven culture. When used efficiently, email is a useful tool for business, and it may be used to handle project or vacation delays by using

automated responses[8].Email is a powerful and safe internet communication tool that businesses and professionals rely heavily on. Email management and tracking is an issue due to its extensive usage, as it takes up a significant amount of users' time. A lot of emails have a set format and call for straightforward answers. Interestingly a study suggested an email response system for frequently asked questions that uses categorization to produce dynamic answers for questions that are repeated.[9].

In order to improve user experience with automated email answers, this study presents a novel technique that combines accurate classification with timely scheduling. For precise email reply predictions, the suggested system makes use of both supervised and unsupervised learning modules. Ensuring the production of accurate, properly structured, and timestamped emails serves as a driving force behind the creation of an effective email reply system that prioritizes engineering features[10].

II. RELATED WORKS

[1]. E-Mail Assistant – Automation of E-Mail Handling and Management using Robotic Process Automation

This paper talks about the problems with email management and how robotic process automation (RPA) is changing company procedures. In order to combat email overload, it presents the application of Case Based Reasoning (CBR) and outlines several methods for email classification. It also looks at how RPA may be used to

manage email traffic and highlights how it can save costs associated with human resources. The section also discusses RPA's research component, advantages, and uses in data-intensive activities and scientific computing.

[2]. Robotic Process Automation: Lessons Learned from Case Studies

The importance of digital tools—in particular, Robotic Process Automation, or RPA—in improving corporate processes is emphasized in this introduction. RPA is a type of lightweight IT that reduces costs and increases productivity by using trained robots to perform manual jobs. The goal of the article is to examine effective RPA implementation by integrating business and academic viewpoints in 10 case studies. The history of RPA, a comparison between BPMS and RPA, the function of Process Mining, and a thorough case study analysis are covered in the following parts.

[3]. A Smart Email Client Prototype for Effective Reuse of Past Replies

The literature emphasizes machine learning techniques while examining different methods for automated email responses. For quicker template retrieval, studies by Sneiders, Kosseim et al., and others emphasize techniques using template-based responses, query-response pair association, and indexing concepts. Kannan et al.'s unique solution, Smart Reply, uses semantic clustering to anticipate brief response recommendations and is deployed in Gmail. In their proactive aid proposal, Van Gysel et al. recommend email attachments according to the message context. Furthermore, the significance of tasks and intent in email messages is determined for better automated classification. Even while rule-based solutions are common, current research notes that creating a solid knowledge base can be difficult, particularly in recently developed systems. It is acknowledged that users want control over automated processes, which highlights the need for solutions that are easy to use.

[4]. Customer Support Emails by RPA Methodology

Operations are becoming automated systems as a result of the use of Robotic Process Automation (RPA) in corporate processes. RPA leverages software robots that are powered by AI and predetermined algorithms. The emphasis on automated email systems brings with it difficulties such as spam emails and massive amounts, which call for effective labeling and classification. In order to tackle email overload,

the study suggests using Case Based Reasoning (CBR), demonstrating how well it maps inquiries based on previous examples.

Numerous methods for email categorization, spam identification, and automation are examined, with a focus on the function of data mining and natural language processing. With the integration of AI, ML, and NLP, intelligent email systems seek to improve user interfaces. The study emphasizes the value of voice-based virtual personal assistants, adaptive information management (AIM) services, email summarization, task classification, scheduling, and prioritization.

[5]. Method of Robotic Process Automation in Invoice Processing and Mailing: Implementation

Given the formal complications brought on by a large number of invoices, Surendhiran Tamilalagan highlights the difficulties that firms encounter when managing their bills. Invoice process management is referred to as a Process Document in the context of Robotic Process Automation (RPA), which shows how business processes can be precisely captured by automation. Sagar Sahu, Sania Salwekar, Atharva Pandit, and Manoj Patil present an automotive application that uses robotic process automation (RPA) to change the way finance operations handles invoices. RPA Data Bot has the ability to improve productivity and lower errors in finance and accounting operations by automating data input, error reconciliation, and decision-making duties associated with processing invoices.

[6]. Customized Automated Email Response Bot using Mac

Understanding email communication techniques, user activities (such as Flow, Triage, Task Management, and Archiving), and user categorization based on historical data or priority were the main points of emphasis. There are two types of automated text summarization: abstractive and extractive. The suggested approach favors extractive summarization, utilizing the TextRank algorithm in particular. The system incorporates RPA technology [5], following the trend of businesses utilizing AI and RPA together to increase productivity and flexibility.

III. SYSTEM ANALYSIS

1) EXISTING SYSTEM

Robotic process automation (RPA) solutions like as Automation Anywhere, Blue Prism, Win Automation,

UiPath, and Automation Anywhere are excellent at automating many business processes, and one area where they shine is email automation. Well-known for its broad application, UiPath excels in accelerating email-related tasks like sending and receiving messages and offers flexibility in automating a variety of business processes. Automation Anywhere is an all-inclusive RPA platform that specializes in email-related activities and provides scalable and flexible solutions. Blue Prism's enterprise-grade platform places a high priority on automation flexibility. Email automation may be easily integrated with its range of customizable functions.

One desktop automation solution that excels at handling email folders and processing attachments is called Win Automation. While selecting an RPA solution for email automation, it is crucial to consider the specific needs of the use case. The flexibility of the selected technology to grow with changing automation requirements should also be taken into account, particularly as email workflows get more complex over time.

2) PROPOSED DESIGN OF THE E-MAIL ASSISTANT

The E-Mail Assistant would have the following functionalities and features.

- A secured robot that could login to the system when allowed to do so.
- Identify the senders without manually checking on devices.
- Save the user's time by not opening the mails from advertising agencies/companies.
- Increase the user's convenience by classifying the mails into various categories.
- Download the attachments if any automatically without human intervention.
- Classify the downloaded attachments into two folders – Useful and Not Useful.
- Inside the Useful folder, further classify them into various folders like – bills, resume, offer letter, invoices, etc.

The system architecture of the suggested email assistant is shown in Fig. 1. It also displays the sequence in which the bot completed each task. The Windows UiPath tool is used in the design and implementation of the bot.

An RPA platform for high-scale, end-to-end automation is called UiPath. This product provides businesses with solutions to automate office monotonous operations in order to quickly improve their business. It can be used to automate repetitive and bothersome chores by utilizing a variety of services [8]. These tools consist of three primary parts: an

email processor engine, an analysis module, and a native email application server. Its autonomous and integrated functionalities are what make it function so well.

1) Native Email Application Server: It accepts incoming emails from the outside world into a public inbox. We call these emails "email loggers."

2) Email Processor Engine: This engine consists of the Segregated Mail Router, RPA Algorithm, Email Logger, Ingestion Layer, and Web Scraping. Each of these modules functions in concert with the others to achieve the goal of email separation.

3) Email Logger: It keeps track of every email along with a ton of information, like the batch ID, sender, recipient, and timestamp.

The emails that are contained in the email logger are comprised of three parts: the header, the topic, and the body, or content. This layer is known as the ingestion layer. Every email has a different ID for it.

4) Web scraping: This method extracts email body text while keeping the unique ID intact. It adds the additional email-related data appendix. This data format's contained information is accepted and ready for analysis in the analyzing module.

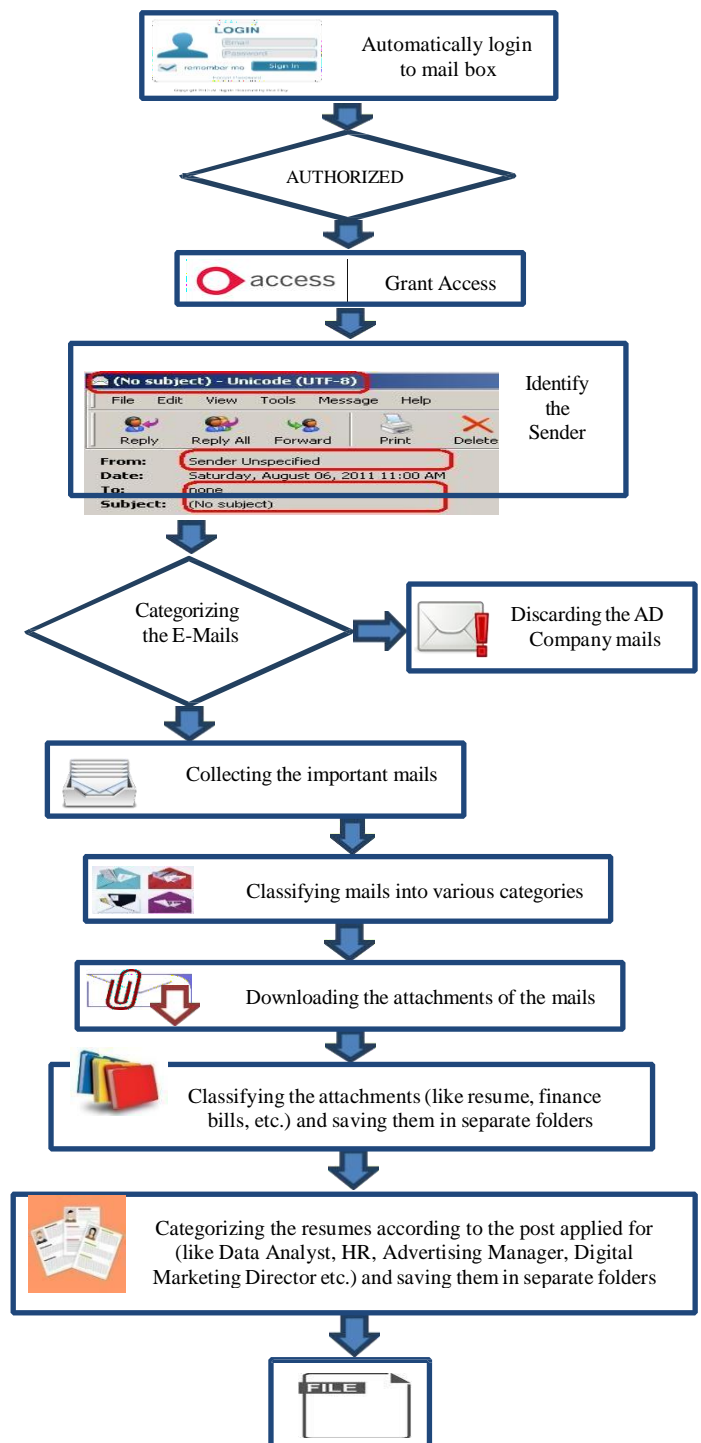
5) Data Enrichment Module: This module optimizes email size by removing features from emails that are noisy or unnecessary.

6) RPA Algorithm: This is a protocol that works with email providers to assign email class label IDs to each email in order to direct it in the appropriate direction.

7) Segregated Mail Router, routes the autoreply emails to the appropriate email addresses. This system divides incoming emails into local folders based on conversation topics so that additional action can be taken. Additionally, it has a prediction module that selects features and represents the email content in order to forecast potential responses to each email individually using the permitted response cluster

8) There are many email activities available in UiPath software and each one is based on the different protocols to be used like Internet Message Access Protocol (IMAP), Post Office Protocol version 3 (POP3), Simple Mail Transfer Protocol (SMTP), Exchange, and Outlook. In the proposed bot, IMAP is used. In the UiPath, the "GetIMAPMailMessages" mail message activity is configured first. It is configured with the Mail Folder name, Port number, Server name, email-id, and password.

Generally, the mail folder name is "Inbox". If the robot is being configured with a Gmail account, then port number 993 is used in most cases. For a Gmail account, the server's name is "imap.gmail.com". Email-id and password can be provided through the window's credential manager [30]. It is also possible to set the number of emails which is executed by the robot in one go. The output argument of this "GetIMAPMailMessages" mail messages activity is a List<MailMessage>. This output can be stored in any variable, say "output_mail". Further, the list of unwanted mail senders which are already known can be provided through an excel sheet, so the mails received from those will be considered not useful.



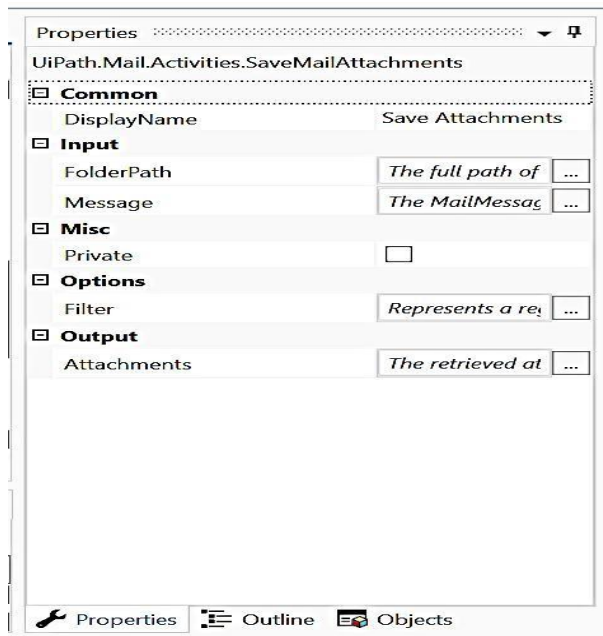
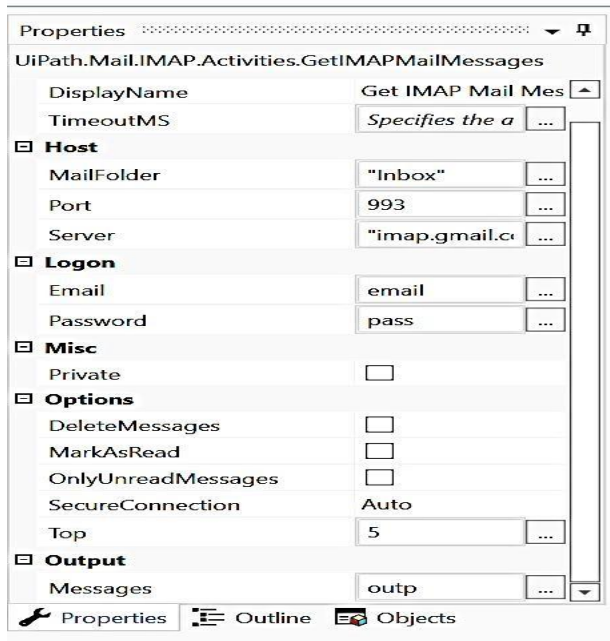


Fig. 2. Configuration of "GetIMAPMailMessages" activity

IV. WORKFLOW EXECUTION

The UiPath version-2022.10.4 (free edition) was installed on the Windows 10 operating system running on the machine with 64-bit Intel(R) Core(TM) i5-8250U CPU @ 1.60GHz and 8 GB RAM. The objective of testing is to demonstrate that the bot is working properly with the functionalities specified

Parameters	Description
TimeoutMS	The amount of time (in milliseconds) activity should be allowed to wait before giving an error message, in case of any delay.
MailFolder	The folder of Gmail should be the working directory for the bot.
Port	The port number on which the robot should receive the emails.
Server	The server's name from which the emails should be extracted ("imap.gmail.com").
Email	The email id of the user can be given at the run time through Window's Credential Manager.
Password	The password for the email id of the user can be given at the run time through Window's Credential manager.
Top	The number of emails that should be taken into work (starting from the top).
Messages	The variable in which all the messages are stored.

TABLE 1. PARAMETERS USED IN THE ACTIVITY

The automation is created like the user wants to automate the process of filtering out unwanted emails, or spam, from their inbox across multiple devices. The initial step establishes a foundation with the creation of a "UnwantedMailAddresses" array for storing undesirable email addresses.

The workflow, utilizing UiPath's "Get Email by Id" activity, ensures targeted retrieval of specific emails, with a focus on systematic analysis using the "For Each" activity. The subsequent steps involve spam checks and removal, culminating in an automated, efficient system for handling unwanted emails in Outlook 365. Trigger setup in UiPath

Conclusion and Future Work

In conclusion, UiPath's RPA-driven Email Automation system excels in streamlining tasks like reading, categorizing, and spam detection. The versatile UiPath robot enhances user experience by securely logging in, categorizing emails, and intelligently organizing attachments. Incorporating spam detection with machine learning and OCR marks a pivotal achievement. Despite successful automation, ongoing RPA maintenance challenges persist.

Future work involves optimizing maintenance, exploring advanced machine learning for spam detection, and adapting to evolving interfaces. Overall, this project establishes a robust foundation for intelligent email automation, promising continued advancements in automated workflows

REFERENCES

- [1] C. L  thje and F. Thiele, "Communication floods--Emails in scholarly communication," *SCM Stud. Commun. Media*, vol. 9, no. 3, pp. 367–393, 2020.
- [2] N. S. Scott et al., "CPC's 50th Anniversary: celebrating 50 years of open-source software in computational physics," *Comput. Phys. Commun.*, vol. 252, p. 107269, 2020.
- [3] E. Milne, *Email and the Everyday: Stories of Disclosure, Trust, and Digital Labor*. MIT Press, 2021.
- [4] G. Abromovich, "Study: Contrary To Popular Belief, Email Is Alive And Well," 2018. .
- [5] S. Chakravarthy, A. Venkatachalam, and A. Telang, "A graph- based approach for multi-folder email classification," in *Proceedings - IEEE International Conference on Data Mining, ICDM*, Dec. 2010, pp. 78–87, doi: 10.1109/ICDM.2010.55.
- [6] N. Jha, D. Prashar, and A. Nagpal, "Combining artificial intelligence with robotic process automation—an intelligent automation approach," in *Deep Learning and Big Data for Intelligent Transportation*, Springer, 2021, pp. 245–264.
- [7] M. Ahmet Unal and O. Bolukbas, "The Acquirements of Digitalization with RPA (Robotic Process Automation) Technology in the Vakif Participation Bank," in *2021 The 4th International Conference on Information Science and Systems*, 2021, pp. 68–73.
- [8] "Robotic Process Automation | UiPath." <https://www.uipath.com/> (accessed Mar. 29, 2021).
- [9] J. Wewerka and M. Reichert, "Robotic Process Automation -- A Systematic Literature Review and Assessment Framework," *arXiv*, Dec. 2020.
- [10] M. Lacity and L. Willcocks, "Becoming Strategic with Intelligent Automation," *MIS Q. Exec.*, vol. 10, no. 2, pp. 1–14, 2021.
- [11] I. Alsmadi and I. Alhami, "Clustering and classification of email contents," *J. King Saud Univ. - Comput. Inf. Sci.*, vol. 27, no. 1, pp. 46–57, Jan. 2015, doi: 10.1016/j.jksuci.2014.03.014.
- [12] M. A. Naeem, I. W. S. Linggawa, A. A. Mughal, C. Lutteroth, and G. Weber, "A Smart Email Client Prototype for Effective Reuse of Past Replies," *IEEE Access*, vol. 6, pp. 69453–69471, 2018.
- [13] S. Park, A. X. Zhang, L. S. Murray, and D. R. Karger, "Opportunities for automating email processing: A need-finding study," in *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, 2019, pp. 1–12.
- [14] M. Peng, B. Gao, J. Zhu, J. Huang, M. Yuan, and F. Li, "High quality information extraction and query-oriented summarization for automatic query-reply in social network," *Expert Syst. Appl.*, vol. 44, pp. 92–101, 2016.