REIMBURSEMENT BOT

A MINI-PROJECT REPORT

Submitted By

BHARGAVI.K 2116220701044

in partial fulfilment for the course

OAI1903 - INTRODUCTION TO ROBOTIC PROCESS AUTOMATION

for the degree of

BACHELOR OF ENGINEERING

IN

COMPUTER SCIENCE AND ENGINEERING

RAJALAKSHMI ENGINEERING COLLEGE
An AUTONOMOUS Institution
Affiliated to ANNA UNIVERSITY, Chennai

RAJALAKSHMI ENGINEERING COLLEGE

AUTONOMOUS, CHENNAI 602105

NOV/DEC,2024

RAJALAKSHMI ENGINEERING COLLEGE

CHENNAI – 602105

BONAFIDE CERTIFICATE

Certified that this project report **"REIMBURSEMENT BOT"** is the Bonafide work of **"BHARGAVI K (220701044)"** who carried out the project work for the subject OAI1903-Introduction to Robotic Process Automation under my supervision.

Mrs. J. Jinu Sophia, M.E., (Ph.D)

SUPERVISOR

Assistant Professor (SG)

Department of Computer Science and Engineering
Rajalakshmi Engineering College, Thandalam

Chennai - 602105

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

INTERNAL EXAMINER **EXTERNAL EXAMINER**

ACKNOWLEDGEMENT

Initially we thank the Almighty for being with us through every walk of our life and showering his blessings through the endeavour to put forth this report. Our sincere thanks to our Chairman **Thiru. S. Meganathan, B.E., F.I.E.,** our Vice Chairman **Mr. M. Abhay Shankar, B.E., M.S.,** and our respected Chairperson **Dr. (Mrs.) Thangam Meganathan, M.A., M.Phil., Ph.D.,** for providing us with the requisite infrastructure and sincere endeavouring in educating us in their premier institution.

Our sincere thanks to **Dr. S. N. Murugesan, M.E., Ph.D.,** our beloved Principal for his kind support and facilities provided to complete our work in time. We express our sincere thanks to **Dr. P. Kumar, M.E., Ph.D.,** Professor and Head of the Department of Computer Science and Engineering for his guidance and encouragement throughout the project work. We convey our sincere and deepest gratitude to our internal guides, **Mrs. J. Jinu Sophia, M.E., Ph.D.,** Assistant Professor (SG) Department of Computer Science and Engineering for their valuable guidance throughout the course of the project. We are very glad to thank our Project Coordinator Professor, **Dr. N. Duraimurugan, M.E., Ph.D.,** Associate Professor and **Mr. B. Bhuvaneswaran, M.E.,** Assistant Professor (SG), Department of Computer Science and Engineering for their useful tips during our review to build our project.

BHARGAVI K - 2116220701044

ABSTRACT

The Reimbursement Bot, developed using UiPath Studio, is a robust Robotic Process Automation (RPA) tool designed to revolutionize the traditionally manual and error-prone process of expense reporting in corporate environments. By automating the extraction of essential data, such as accommodation, travel, and food expenses, from PDF invoices, the bot streamlines reimbursement management and reduces the time and effort required for data entry. It leverages OCR (Optical Character Recognition) technology to accurately read and extract information, then organizes it into a structured Excel report, providing a clear and organized overview of expenses. This automation enhances accuracy, speeds up processing times, and ensures consistent data handling, reducing the risk of errors and discrepancies. As a result, the bot not only improves efficiency but also frees up valuable resources for strategic decision-making, while offering significant cost savings by eliminating the need for manual data entry and verification. The Reimbursement Bot ultimately transforms the expense reporting process, offering a scalable, reliable, and efficient solution for organizations

.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF ABBREVIATIONS

Abbreviation	Full Form
API	Application Programming Interface
FFmpeg	Fast Forward Moving Picture Experts Group
ML	Machine Learning
UiPath	A Robotic Process Automation (RPA) tool
OCR	Optical Character Recognition
XML	Extensible Markup Language
JSON	JavaScript Object Notation
API Key	Application Programming Interface Key
CSV	Comma Separated Values

CHAPTER-1
INTRODUCTION

1.1 INTRODUCTION

The Reimbursement Bot, developed using UiPath Studio, is an automated solution designed to streamline the traditionally manual and time-consuming process of expense reporting in corporate environments. Managing employee reimbursements for expenses such as accommodation, travel, and food often involves tedious tasks like reviewing invoices, verifying amounts, and entering data into spreadsheets. These manual processes are prone to errors and delays, which can lead to inefficiencies and discrepancies in financial reporting.

By leveraging Robotic Process Automation (RPA) and Optical Character Recognition (OCR) technology, the bot automates the extraction of key data from PDF invoices, eliminating the need for manual intervention. The extracted data is then organized into a structured Excel report, ensuring accuracy and improving efficiency. This automation not only reduces human errors and speeds up the reimbursement process but also provides a scalable, reliable solution for managing employee expenses in a more efficient and transparent manner.

1.2 OBJECTIVES

The Reimbursement Bot's primary objectives are to enhance the efficiency, accuracy, and scalability of the reimbursement workflow.

1 **Automate Data Extraction**: To extract key information such as accommodation, travel, and food expenses from invoice PDFs using UiPath's OCR capabilities.
2 **Enhance Efficiency**: Minimize manual effort in processing invoices by automating repetitive and time-consuming tasks.
3 **Ensure Accuracy**: Validate the extracted data to reduce errors and improve the reliability of the reimbursement process.
4 **Generate Consolidated Reports**: Create an organized Excel sheet summarizing all reimbursable expenses for easy review and processing.
5 **Improve Scalability**: Develop a flexible system capable of handling a high volume of diverse invoice formats and data.
6 **Support Decision-Making**: Provide users with clear and structured expense data to facilitate faster approvals and reimbursements.

1.3 EXISTING SYSTEM

1. Manual Effort: The current reimbursement process relies heavily on manual review and data entry, where individuals extract expense details such as accommodation, travel, and food costs from invoices. This is labour-intensive and prone to errors.

2. Time Consumption: Processing a large volume of invoices is time-consuming, leading to delays in approvals and reimbursements. Employees must allocate significant effort to repetitive tasks.

3. Inaccuracy Risks: Human errors in extracting and inputting data can result in inaccuracies, potentially leading to disputes, incorrect reimbursements, or financial discrepancies.

4. Limited Tools and Operational Inefficiency:

Existing invoice management tools often lack advanced automation capabilities, requiring significant human intervention to handle diverse invoice formats. This reliance on manual processes, coupled with the limitations of basic software, leads to increased operational costs, reduced efficiency, and delays in the reimbursement workflow.

1.4 PROPOSED SYSTEM

The proposed "Reimbursement Bot" utilizes UiPath Studio to automate expense data extraction from invoices, ensuring accurate, efficient, and scalable reimbursement processing. By consolidating extracted data into Excel reports, the system addresses inefficiencies in traditional workflows.

1 **Automation of Invoice Processing:** Utilizes UiPath Studio and OCR to extract expense details from invoices, reducing manual effort and time.

2 **Accurate Data Extraction:** Ensures precise extraction of relevant fields with validation mechanisms, handling diverse formats effectively.

3 **Dynamic Report Generation**: Compiles extracted data into structured Excel reports for easy review and consolidated insights.

4 **Scalability and Flexibility**: Handles high invoice volumes and adapts to diverse workflows, suitable for organizations of all sizes.

5 **Improved Efficiency and Cost Reduction**: Streamlines reimbursements, reduces errors, accelerates approvals, and lowers operational costs.

CHAPTER-2

LITERATURE REVIEW

2.1 GENERAL

The "Reimbursement Bot" leverages automation technologies such as RPA, OCR, and AI to streamline the reimbursement process. By automating the extraction and validation of expense data from invoices, the bot reduces errors and accelerates processing. The literature highlights the advantages of these technologies in managing invoices, categorizing expenses, and ensuring compliance with organizational policies, ultimately improving efficiency and accuracy in reimbursement workflows.

1. Automating the Invoice Process

This article outlines how UiPath's Robotic Process Automation (RPA) and Optical Character Recognition (OCR) technologies are used to automate invoice processing, significantly reducing the manual effort involved in extracting and validating data. The use of UiPath’s RPA tools enables efficient handling of high-volume invoices by automating tasks such as data entry, categorization, and validation across multiple invoice formats. OCR technology helps extract structured data, even from scanned or image-based invoices, making it an ideal solution for automating the reimbursement process and ensuring accuracy and efficiency.

2. Expense Policy

The IEEE Expense Policy defines the procedures for handling travel and other reimbursable expenses within the organization. It includes detailed guidelines on acceptable expense categories, such as accommodation, travel, meals, and necessary documentation required for reimbursement claims. This policy ensures that automated reimbursement systems, such as the Reimbursement Bot, align with IEEE's compliance requirements. By adhering to these standards, the bot ensures that all extracted data is categorized correctly and complies with IEEE's reimbursement protocols, thereby supporting accurate and consistent expense processing.

3. Reimbursement Procedures for Travel and Other Expenses

This document provides step-by-step instructions for submitting and processing reimbursement claims for travel and other expenses. It stresses the importance of proper documentation, clear categorization of expenses, and timely submissions. The guidelines also indicate the need for systems that can handle diverse invoice formats and produce accurate reports for approval. These procedures serve as the foundation for the Reimbursement Bot, ensuring that the automated solution adheres to IEEE's reimbursement guidelines while efficiently managing expense claims.

The literature reviewed provides valuable insights into the automation of business processes, particularly in the areas of invoice processing, expense management, and customer support. The integration of UiPath's RPA and OCR technologies, combined with the IEEE expense policies and customer service automation principles, forms the foundation for the "Reimbursement Bot." This bot automates the extraction, validation, and reporting of expenses, ensuring compliance, reducing human error, and improving overall efficiency in the reimbursement process.

CHAPTER-3

SYSTEM DESIGN

3.1 SYSTEM FLOW DIAGRAM:

The system flow of the "Reimbursement Bot" begins with the user uploading invoice PDFs for processing. The bot then uses OCR to preprocess the invoice, converting the text into machine-readable data. It extracts relevant expense details, such as accommodation, travel, and food costs, from the invoice. After extraction, the bot validates the data to ensure it complies with predefined expense policies, such as correct categorization and value ranges. The validated data is then compiled into a structured Excel report, which is saved or sent to the user for review

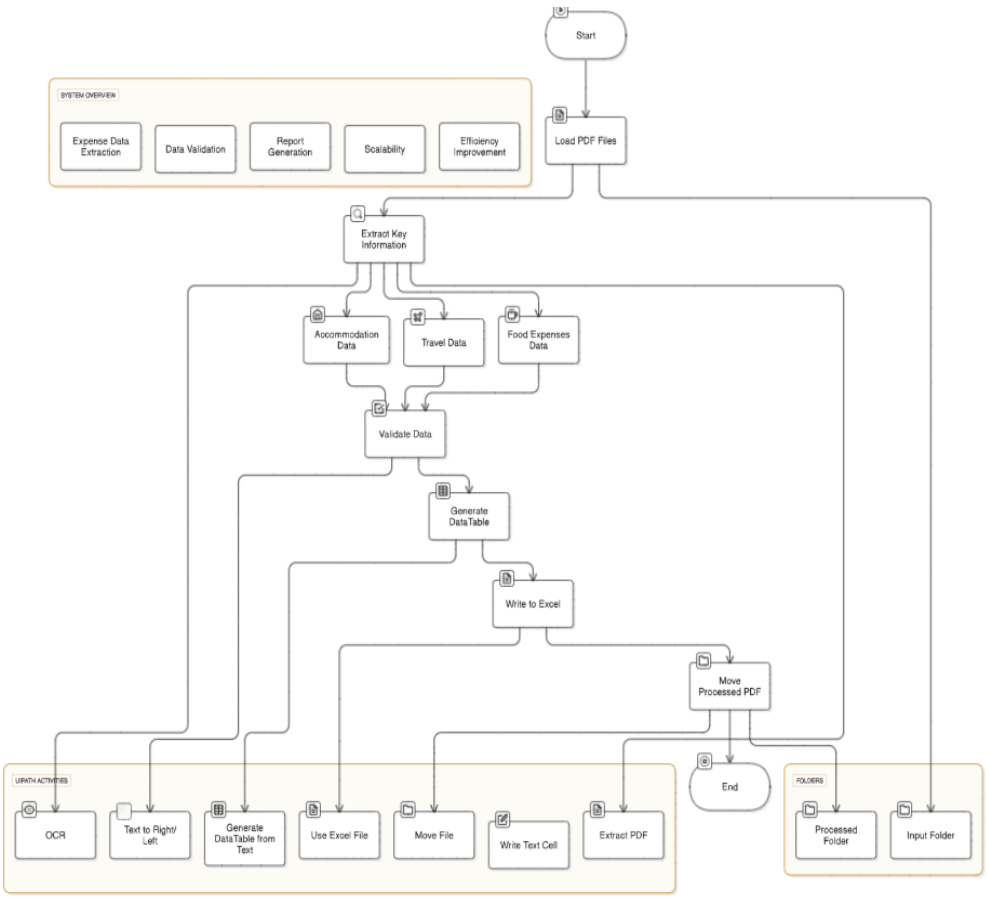


FIG 1

3.2 ARCHITECTURE DIAGRAM

1 **User Interaction & Input Folder**: The user uploads invoice PDFs, which are stored in the Input Folder for processing by the bot.

2 **Reimbursement Bot** (Core): The bot orchestrates the entire process, from OCR text extraction to data validation and report generation.

3 **Data Extraction & Validation**: The bot uses OCR to extract data, then categorizes expenses and validates it against predefined policies.

4 **Report Generation & File Handling**: Validated data is written into an Excel file, saved in the Processed Folder, and the processed invoices are moved to archive.

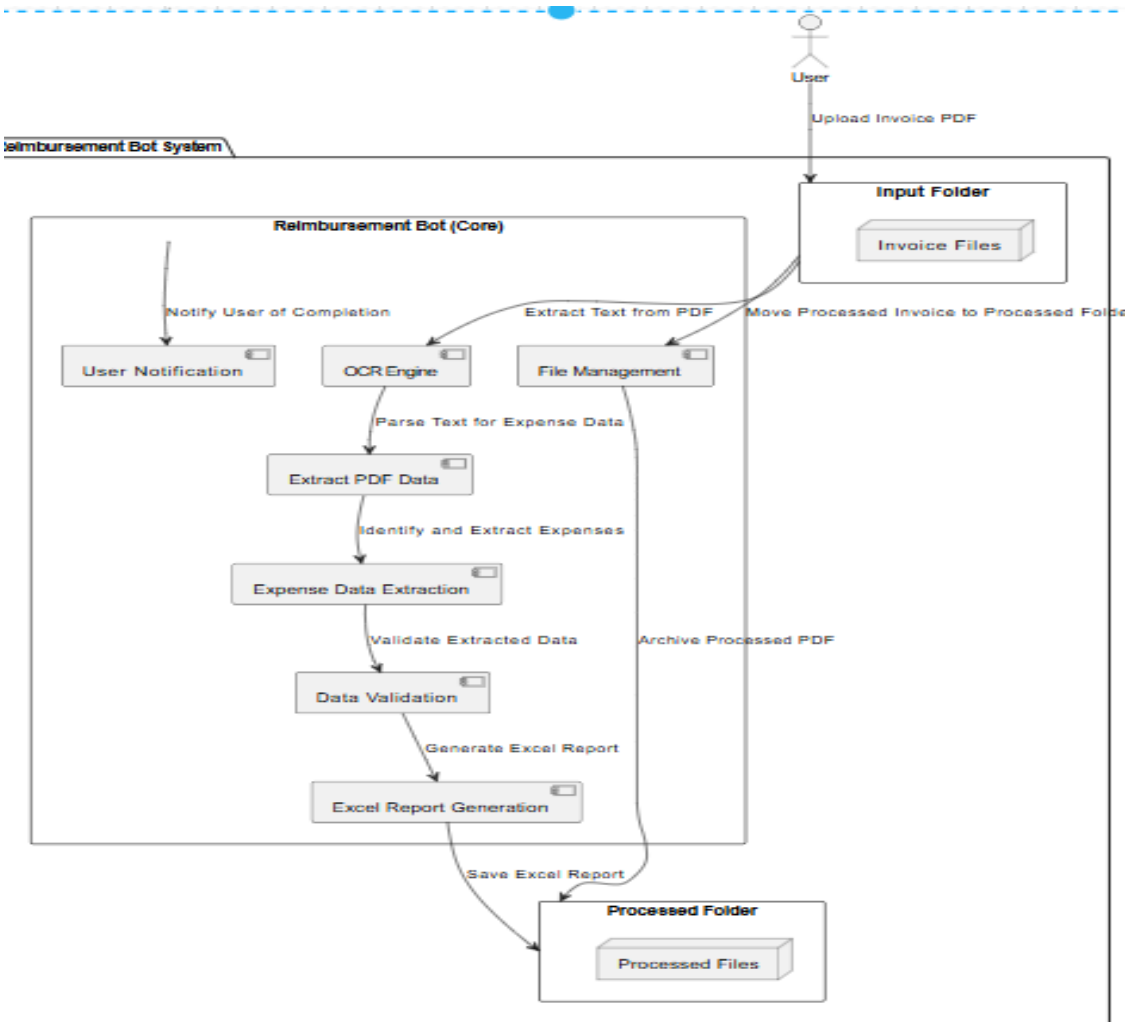


Fig 2

3.3 SEQUENCE DIAGRAM

1 User uploads invoice PDF to the Input Folder, which notifies the Reimbursement Bot.

2 The Reimbursement Bot sends the invoice to OCR to extract text, which is processed by Extract PDF for data extraction.

3 The extracted data is passed to Expense Data Extraction and validated by Data Validation to ensure compliance with expense policies.

4 The validated data is organized into a table by Generate Data Table, then written into an Excel report by Write to Excel.

5 The final report is saved in the Processed Folder, and the Reimbursement Bot notifies the User that the process is complete.

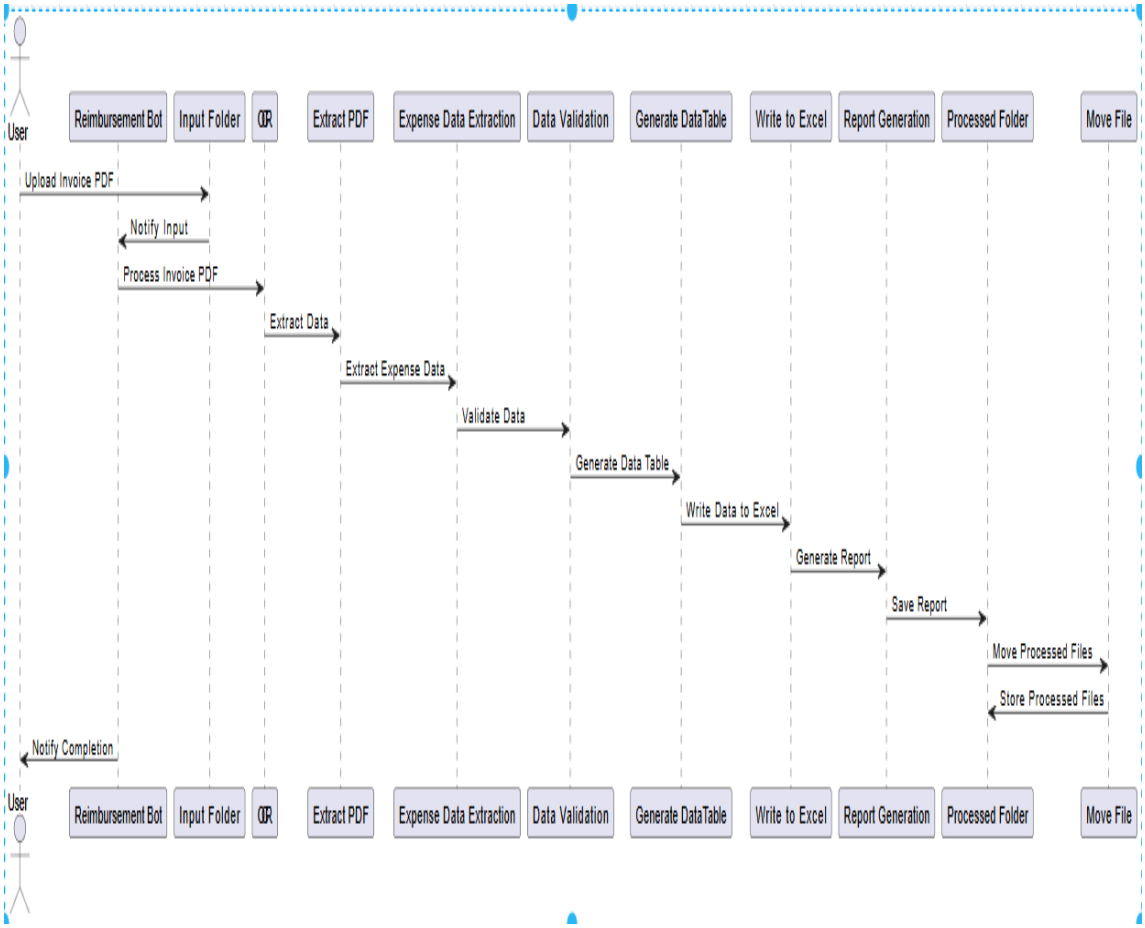


Fig 3

CHAPTER-4
PROJECT DESCRIPTION

4.1 METHODOLOGIES

The "Reimbursement Bot" employs Robotic Process Automation (RPA) and Optical Character Recognition (OCR) to automate the entire invoice processing workflow. The bot begins by extracting text from invoices using OCR, followed by parsing and categorizing the extracted data into predefined expense categories like accommodation, travel, and food. The data is then validated to ensure it meets company policies. After validation, the bot structures the data into a table format and generates an Excel report. The processed invoice PDFs are archived, and the final report is stored in the Processed Folder. The system is designed to be scalable, ensuring efficient handling of large volumes of invoices with minimal human intervention, improving accuracy and reducing operational costs.

4.2 MODULES

Invoice PDF Extraction

- **Tool**: OCR (Optical Character Recognition)
- **Process**: The first module focuses on extracting text from the uploaded invoice PDFs. Using OCR technology, the bot reads the text from scanned or image-based PDFs and converts it into machine-readable content. This step is essential for transforming the unstructured invoice data into a format that can be processed further. OCR is preferred for its high accuracy, versatility, and ability to handle various document layouts. The text extracted serves as the input for subsequent data categorization and validation.

Generate Data Table

- **Tool**: Data Structuring Algorithms
- **Process**: After the data is validated, this module organizes the information into a structured table format, typically using rows and columns, that can be easily written into an Excel file. The table organizes the categorized expense data, such as accommodation costs, travel expenses, and food allowances, making it suitable for further processing or reporting. This step helps in generating a standardized format for the final report, ensuring consistency across all processed invoices.

Excel Report Generation

- **Tool**: Excel Automation (e.g., UiPath Excel Activities)
- **Process**: This module takes the structured data from the "Generate Data Table" step and writes it into an Excel file. It formats the data, applying necessary styles such as bold headers, column alignment, and number formatting to ensure the report is easy to read. The Excel file serves as the final output, summarizing all extracted and validated expense data, and is saved for further review or approval. This module ensures that the reimbursement data is organized and accessible for administrative review.

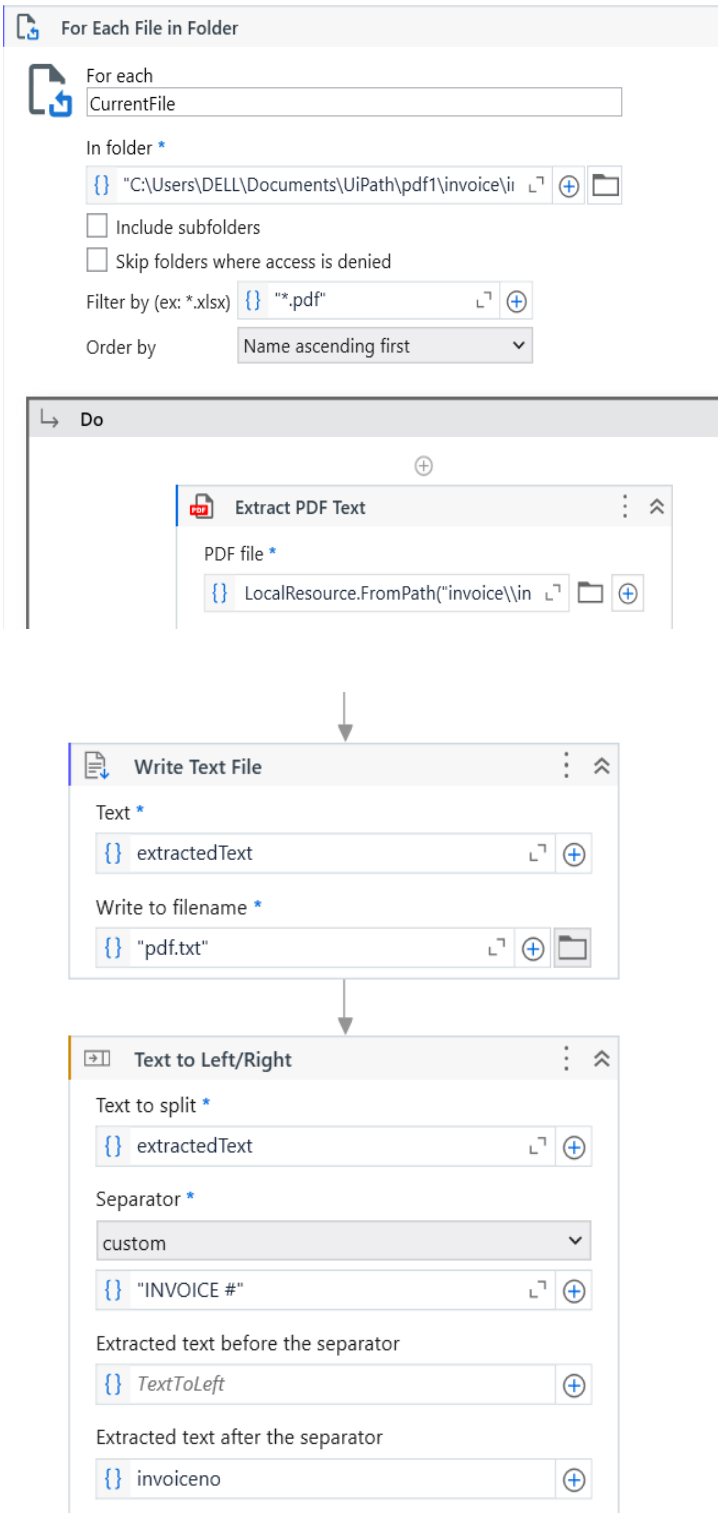
File Management

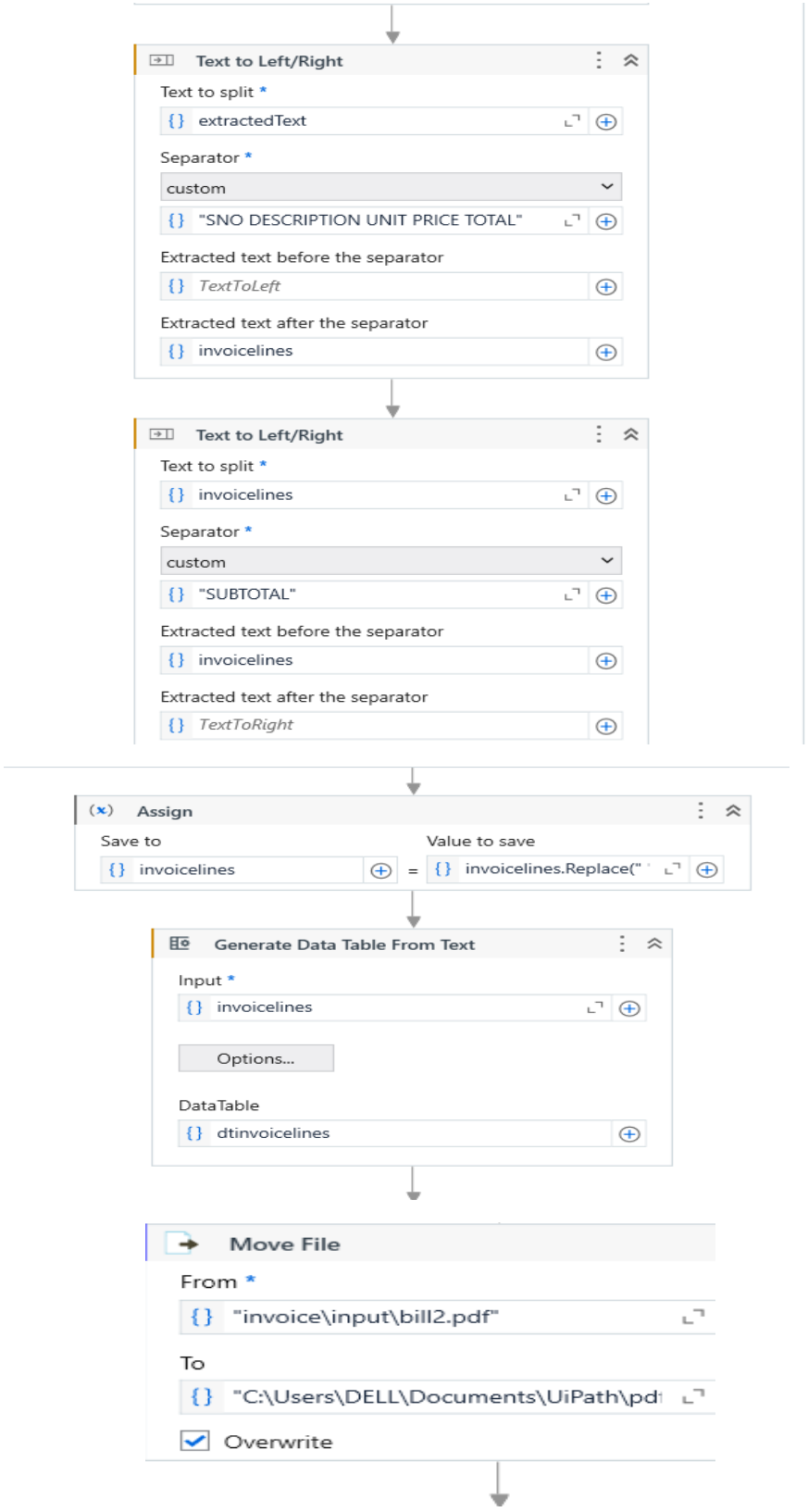
- **Tool**: File System Operations
- **Process**: After processing and generating the report, this module handles file management tasks. It moves the processed invoice PDF from the **Input Folder** to the **Processed Folder** for archiving purposes. Additionally, the generated Excel report is stored in the appropriate location for easy retrieval. The bot ensures that both the invoice and the report are organized and securely stored, providing a traceable history of all processed files.

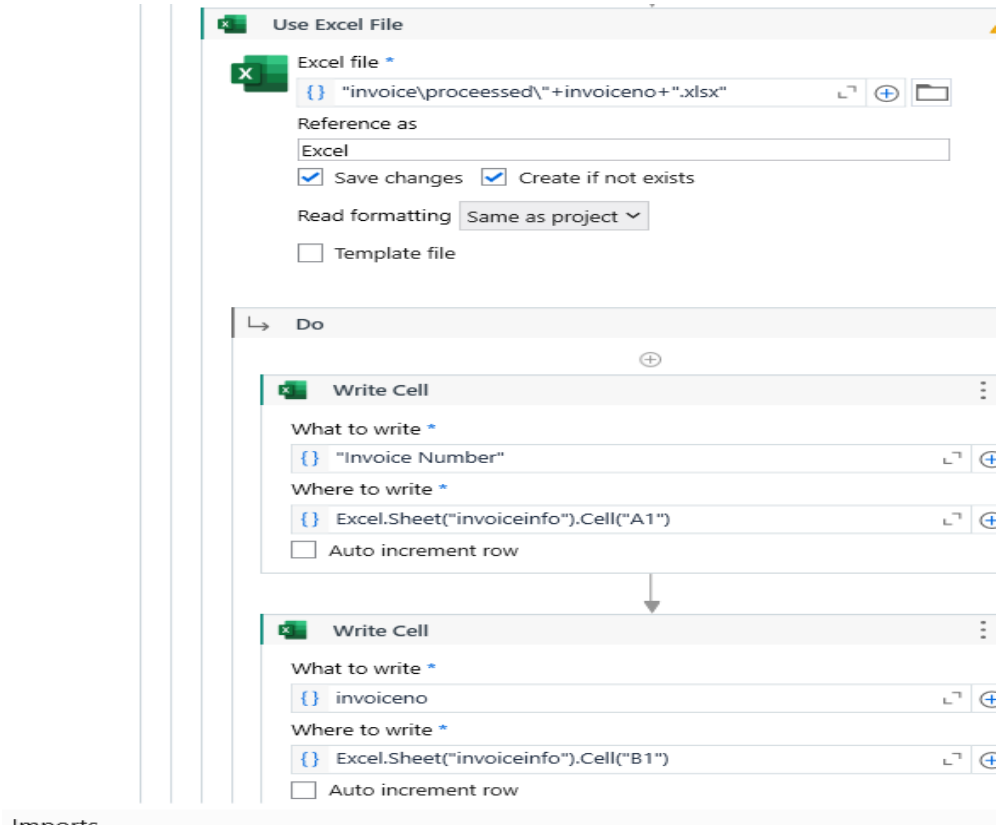
CHAPTER-5

OUTPUT SCREENSHOTS

5.1 IMPLEMENTATION WORKFLOW




Text to Left/Right
Text to split *
extractedText
Separator *
custom
"SNO DESCRIPTION UNIT PRICE TOTAL"
Extracted text before the separator
TextToLeft
Extracted text after the separator
invoicelines
Text to Left/Right
Text to split *
invoicelines
Separator *
custom
"SUBTOTAL"
Extracted text before the separator
invoicelines
Extracted text after the separator
TextToRight
Assign
Save to
Value to save
invoicelines
=
invoicelines.Replace("
Generate Data Table From Text
Input *
invoicelines
Options...
DataTable
dtinvoicelines
Move File
From *
"invoice\input\bill2.pdf"
To
"C:\Users\DELL\Documents\UiPath\pd
Overwrite


Use Excel File
Excel file *
{} "invoice\proceessed\"+invoiceno+".xlsx"
Reference as
Excel
Save changes
Create if not exists
Read formatting
Same as project
Template file
Do
Write Cell
What to write *
{} "Invoice Number"
Where to write *
{} Excel.Sheet("invoiceinfo").Cell("A1")
Auto increment row
Write Cell
What to write *
{} invoiceno
Where to write *
{} Excel.Sheet("invoiceinfo").Cell("B1")
Auto increment row

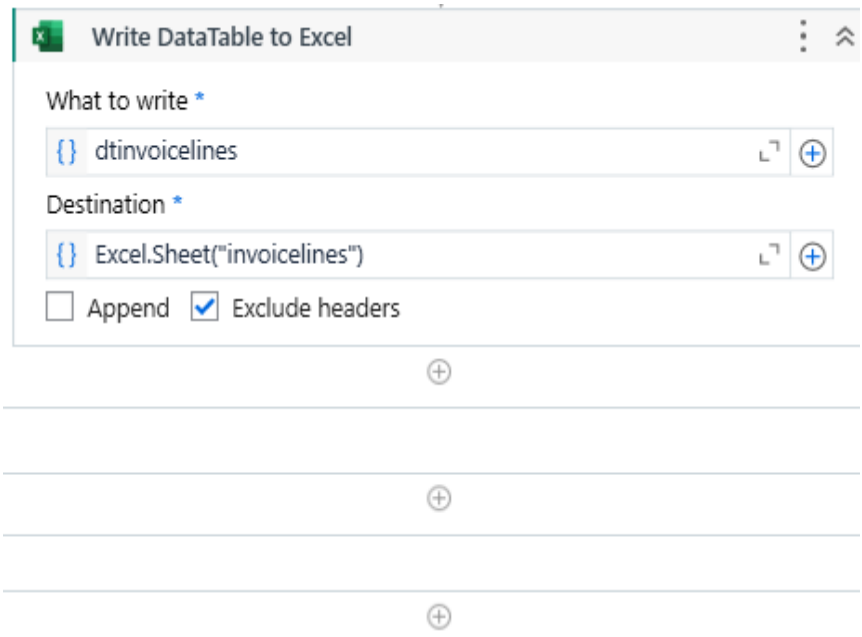

Write DataTable to Excel
What to write *
{} dtinvoicelines
Destination *
{} Excel.Sheet("invoicelines")
Append
Exclude headers

Fig 4

5.2 INPUT FORM

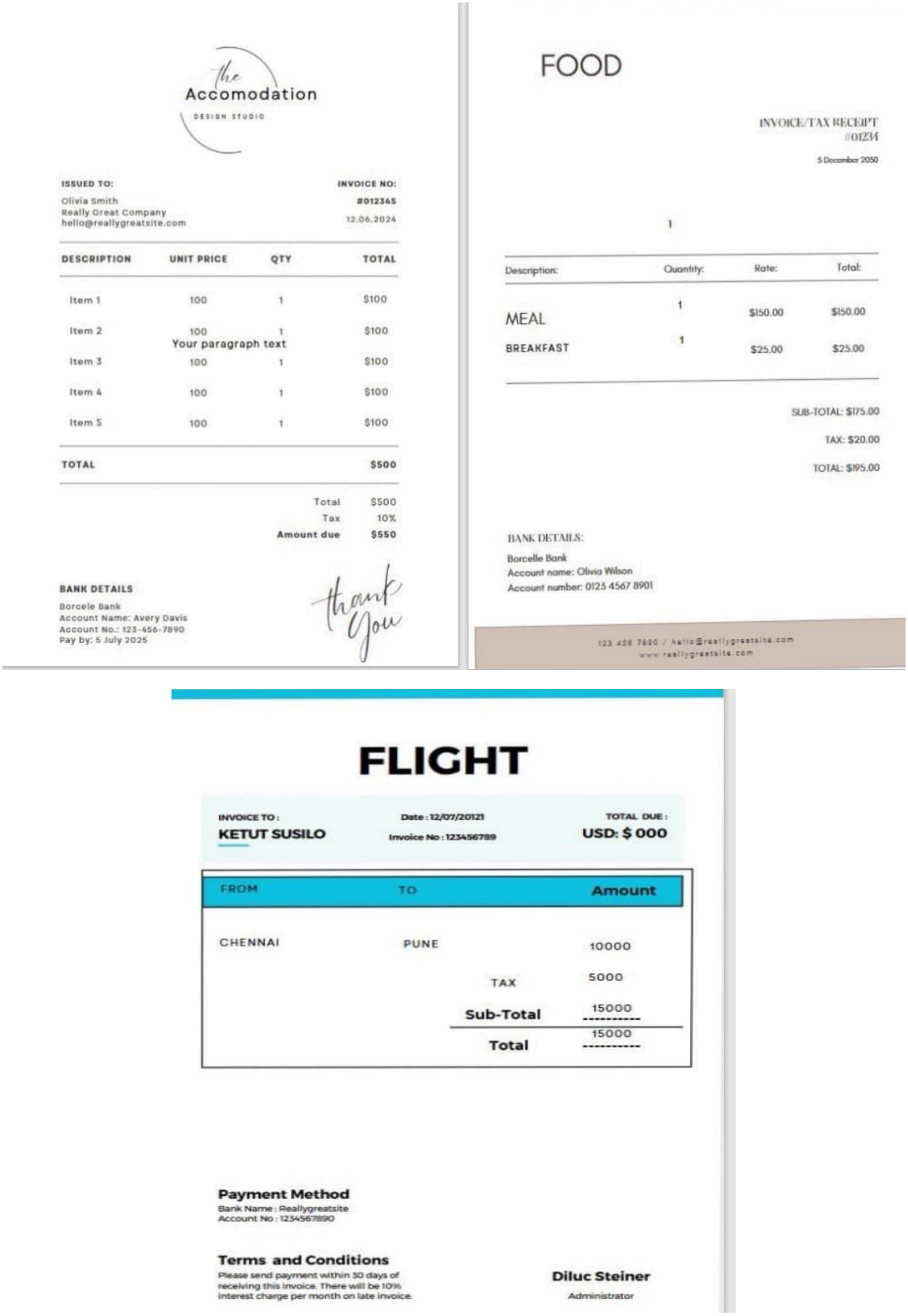
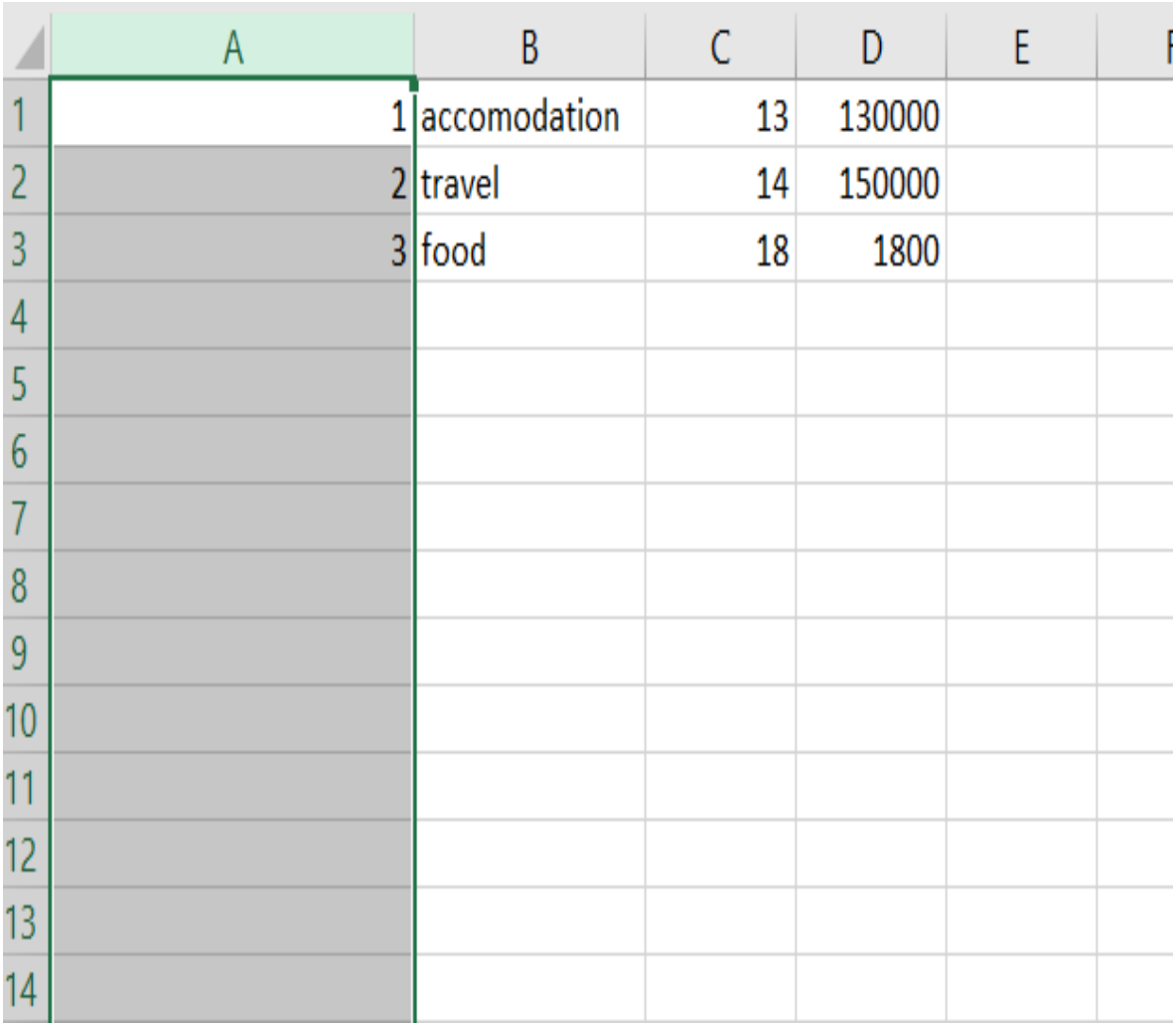


Fig 5

5.3 OUTPUT

The report is generated and given in excel sheet and the input pdf is moved to processed folder



	A	B	C	D	E
1	1	accomodation	13	130000	
2	2	travel	14	150000	
3	3	food	18	1800	
4					
5					
6					
7					
8					
9					
10					
11					
12					
13					
14					

Fig 6

CHAPTER-6

CONCLUSION

GENERAL

The "Reimbursement Bot" effectively automates the invoice processing workflow, transforming a traditionally manual task into an efficient, error-free process. By using Optical Character Recognition (OCR) technology, the bot extracts text from scanned or image-based invoices, making the data machine-readable. Custom parsing logic categorizes expenses such as accommodation, travel, and food, while data validation ensures that the extracted data adheres to company policies, minimizing errors and ensuring compliance.

Once the data is validated, the bot generates a structured Excel report, which organizes the expenses into a format that is easy for administrators to review and approve. This automation eliminates the need for manual data entry, reducing time spent on processing invoices and improving accuracy. The Excel report serves as a reliable record for expense tracking and approval.

Overall, the "Reimbursement Bot" streamlines the reimbursement process, making it more efficient and transparent. The solution is scalable, allowing it to handle large volumes of invoices with minimal human intervention, saving operational costs while ensuring the integrity and consistency of financial reporting. This project demonstrates the potential of automation to simplify and improve administrative tasks in organizations.

APPENDICES

Appendix A: Workflow Design

1. Invoice Upload: User uploads invoice PDFs to the Input Folder, where the bot detects and processes them.
2. OCR Text Extraction: The bot uses OCR to extract text from the uploaded PDF invoices for further processing.
3. Expense Data Extraction & Validation: The bot categorizes and validates the extracted expense data according to predefined policies.
4. Excel Report Generation: The bot organizes the validated data into a structured Excel report for easy review.
5. File Management & User Notification: The bot moves processed invoices to the Processed Folder and notifies the user once the process is complete.

Appendix B: Tools and Technologies Used

1. **UiPath Studio**: Used for designing and automating the entire RPA workflow.
2. **OCR (Optical Character Recognition)**: Extracts text from scanned or image-based invoices.
3. **Microsoft Excel (via UiPath Excel Activities)**: Generates and formats structured Excel reports.
4. **Custom Parsing Logic**: Identifies and categorizes expense data from extracted text.
5. **File System Operations**: Manages files in the **Input Folder** and **Processed Folder**.

Appendix C: System Requirements

1. **Hardware Requirements**:
    - **Processor**: Minimum 2.0 GHz Dual-Core Processor or higher.
    - **RAM**: At least 4 GB of RAM
    - **Storage**: Minimum 10 GB of free disk space for installation and file storage.

2. **Software Requirements**:
    - **Operating System**: Windows 7 or higher (64-bit).
    - **UiPath Studio**: Latest version of UiPath Studio for creating and running automation workflows.
    - **Microsoft Excel**: Microsoft Excel 2016 or higher for report generation (via UiPath Excel Activities).
    - **OCR Engine**: Integration with OCR tools

REFERENCE

[1] https://www.uipath.com/blog/product-and-updates/uipath-at-work-automating-the-invoice-process

[2] https://ieeeaps.org/membership/ieee-expense-policy

[3] https://www.ieee-ras.org/about-ras/volunteer