Process Definition Document (PDD)

Project Name: Automated Invoice Processing Bot using Blue Prism

Date: May 27, 2025

Author: RPA Developer

# 1. Introduction

This document defines the process for an RPA bot built using Blue Prism that automates the extraction of invoice data from email attachments and records the data into an Excel spreadsheet.

# 2. Business Process Overview

The finance department receives daily emails containing invoice PDFs. These invoices need to be recorded in an Excel tracker. This bot will automate the retrieval of these emails, extraction of invoice data, and entry into Excel.

# 3. In-Scope

- Access Outlook inbox

- Filter emails with subject line containing 'Invoice'

- Download PDF attachments

- Extract invoice data using a free PDF parser

- Enter extracted data into an Excel sheet

# 4. Out-of-Scope

- PDF files requiring passwords or licenses

- Manual approval processes

- Integration with actual ERP systems

# 5. Process Steps

1. Open Outlook and read unread emails with subject containing 'Invoice'

2. Download attached PDF files to a local folder

3. Use a free PDF parser to extract: Invoice Number, Vendor, Date, Amount

4. Validate data format and values

5. Append the data into an Excel tracker

# 6. Exception Handling

- If no emails are found, log and end process

- If PDF is unreadable, move it to a failed folder and log

- If required data is missing, log and continue with the next file

# 7. Requirements

- Microsoft Outlook configured and accessible

- Blue Prism installed

- Free PDF parser object (e.g., PDFTextReader VBO)

- Excel application installed

# 8. Assumptions

- All invoice PDFs are text-based and do not require OCR

- PDF structure is consistent across all files

# 9. Sample Output

The Excel sheet will include the following columns:

Invoice Number | Vendor | Invoice Date | Amount