Process Definition Document

*Process Name: Invoice Scraping*

Table of Contents

[**Introduction**](#_fmc2ik42b62t) **1**

[Purpose of the Document](#_8b0nhjcbe7cw) 1

[Objectives](#_soa72miybokv) 1

[Process Key Contact](#_e5eh7vtp3elw) 1

[Minimum Prerequisites for Automation](#_8uc76jjm25ud) 1

[**As-Is Process Description**](#_mdr6kpc5a5r5) **2**

[Process Overview](#_vrc3lxjwb5na) 2

[Applications used in the Process](#_in5ehl2op8tm) 3

[As-Is Process Map](#_3xrlx7nhtlp7) 3

[**To-Be Process Description**](#_gmvdjkbe065o) **4**

[Detailed Process Map](#_nmn90y3pi1ee) 4

[Robot Type](#_cx5ym07ptgjk) 4

[Business Exceptions Handling](#_7u1z1cuc6dh1) 4

[Known Exceptions](#_jr6jw3koor93) 4

[Unknown Exceptions](#_57bwdp6ycy5h) 5

[System Exceptions Handling](#_1os2bz2dwbrb) 5

[**Other Observations**](#_bsbyy5x0t0oj) **5**

[**Additional sources of process documentation**](#_ms60s1nz1shm) **5**



# Introduction

## Purpose of the Document

The Process Definition Document outlines the business process chosen for automation using UiPath Robotic Process Automation (RPA) technology.

The document describes the sequence of steps performed as part of the business process, the conditions and rules of the process prior to automation and how they are envisioned to work after automating it, partly or entirely. This specifications document serves as a base for developers, providing them with the details required for applying robotic process automation to the selected business process.

## Objectives

The process has been selected for RPA as part of the project initiative conducted within Techno Computers Inc., the Finance department.

The objective of this process automation is linked to the project business case and is mainly intended to:

* Deliver faster processing
* Reduce redundant activities
* Improve overall performance and reliability

## Process Key Contact

The specifications document includes concise and complete requirements of the business process and it is built based on the inputs provided by the process Subject Matter Expert (SME)/ Process Owner.

The Process Owner is expected to review it and provide signoff for accuracy and completion of the steps, context, impact and a set of process exceptions. The details are to be included in the table below.

|  |  |  |  |
| --- | --- | --- | --- |
| Role | Name | Contact Details (email & phone number) | Notes |
| Process Owner | Jesse Peterson | petersoninquiries@gmail.com |  |
| Business Analyst | Jesse Peterson | petersoninquiries@gmail.com |  |

## Minimum Prerequisites for Automation

|  |  |
| --- | --- |
| Met (Y/N) | Prerequisites |
|  | A filled in and completed Process Definition Document |
|  | Closure of any open process questions |
|  | Environment set up |
|  | Test Data to support development and testing |
|  | User access and creation of user accounts (licences, permissions, restriction to create accounts for robots) |

# As-Is Process Description

## Process Overview

General information about the process selected for RPA prior to automation.

|  |  |  |
| --- | --- | --- |
| # | Item | Description |
| 1 | Process Full Name | Invoice Scraping Process |
| 2 | Process Area | Personal |
| 3 | Department | Student |
| 4 | Process Short Description  (operation, activity, outcome) | A process that will scrape relevant data from the invoices for further processing. The Robot will read through emails and download the invoices received in the form of an email attachment as a PDF. It will extract specific data and store those values in an Excel spreadsheet and a subset of values will be uploaded to the Orchestrator Queue. And finally, the robot will email the spreadsheet to yourself when finished. |
| 5 | Role(s) required for performing the process | Any |
| 6 | Process schedule and frequency | As needed (recommended End of Day [EOD]) |
| 7 | # of items processed /reference period | 4 invoices |
| 8 | Process execution time | 2-3 minutes/invoice |
| 9 | Peak period(s) | N/A |
| 10 | Transaction Volume During Peak period | N/A |
| 11 | Total # of FTEs supporting this activity | N/A |
| 12 | Expected increase of volume in the next reference period | N/A |
| 13 | Level of exception rate | N/A |
| 14 | Input data | Invoices as an attachment over email |
| 15 | Output data | Order details uploaded to Orchestrator Queue |

\*Add more rows to the table to include relevant data for the automation process. No fields should be left empty. Use “n/a” for the items that don't apply to the selected business process.

## Applications used in the Process

The table includes a comprehensive list of all the applications that are used as part of the process to be automated to perform the given steps in the flow.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| # | Application Name & Version | System Language | Thin/Thick Client | Environment/ Access Method | Comments |
| 1 | Microsoft Outlook Version 2006 | English | Thin | PC |  |
| 2 | Microsoft Excel 2016 | English | Thin | PC |  |
| 3 | Adobe Reader PDF | English | Thin | PC |  |

\*Add more rows to the table to include the complete list of applications.

**-------------Complete the rest of the document and submit along with your final submission.-------------**

## As-Is Process Map

**High Level As-Is Process Map:** This chapter depicts the As-Is business process at a High Level to enable developers to have a high-level understanding of the current process.

Send email with invoice attached

Upload scraped invoice data to UiPath Orchestrator

Write scraped invoice data to Excel file

Scrape invoice header, subtotal, and item details

Downloads Invoices  
form email

**Detailed Process Map:** This chapter depicts the As-Is business process at a detailed view to enable process owners to document their process

**Orchestrator**

**End**

**Invoice**

**PDF**

**6**

**4**

**Excel**

**File**

**5**

**3**

**2**

**1**

|  |  |  |  |
| --- | --- | --- | --- |
| # | Step Action/Description | Screenshot | Remarks |
| 1 | Download Invoice |  | Downloads invoice to a ‘Data’ folder |
| 2 | Load Invoice and Scrape Invoice Number, Customer Name |  | Extracts Invoice Number and Customer Name from file name |
| 3 | Scrape Invoice Item Details and Totals |  | Extracts Item Details and Totals from PDF file contents |
| 4 | Upload Invoice Totals to UiPath Orchestrator |  | Uploads Subtotal, GST, and Total information to an existing Orchestrator Queue |
| 5 | Extract Invoice Number, Date, Item Details, and Totals to Excel file |  | Create new Excel for each invoice |
| 6 | Email Excel file |  | Send email with custom subject line and Invoice Excel file as an attachment |

# To-Be Process Description

## Detailed Process Map

**High Level To-Be Process Map:** This chapter depicts the To-Be automation process at a High Level to enable developers/COE to have a high-level understanding of the to be developed process.

**Detailed Process Map:** This chapter depicts the To-Be automation process at a detailed view to enable developers/COE to see the workflows involved in the RPA solution

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Workflow Name** | **Description** | **Pre-conditions** | **Post-actions** | **Arguments** | **Notes** |
| Main.xml | 1). Downloads invoices from a email inbox based on a specific subject line and those containing attachments.  2). Opens the downloaded invoices then scrapes and filters multiple fields and data tables from the invoices.  3). Writes scrapped invoice data to an Excel file and saves it based on original invoice name.  4). Uploads scrapped fields from the invoice to an Orchestrator Queue.  5). Send the Excel file with the scrapped invoice data to an email address with a custom subject line. | - Users send and receive emails from a gmail accont  - The UiPath Robot is connected to an orchestrator account with the ability to write to an existing Queue | - Sends PDF invoice files as an attachment to a Gmail account  - Uploads scrapped data from PDF invoices to Orchestrator Queue | None |  |
| Email\_Excel\_Sequence.xml | 1). Logs into an email account  2). Sends an email to an account with a custom subject line and the email excel spreadsheet with scraped invoice data as an attachment | - Receives a list of Invoice files to be emailed | - Sends scraped invoice data to email with customer subject line and scraped excel files as an attachment | - in\_Subject  - in\_Attachment |  |
| Read\_Email\_Sequence.xml | 1). logs into an email account and downloads attachments from unread messages with a specific subject line to a local folder, creating this folder if it doesn't exist | - A user has entered their email credentials | - Downloads email attachment to a local ' Data ' folder | None |  |
| Read\_PDF\_Write\_Excel\_Sequence.xml | 1) Iterates through invoice files in a local folder  2). Scrapes the invoice number and customer names from each invoice  3). Calls a workflow to scrape the invoice item details, passing the invoice files as an argument  4). Calls a workflow to upload invoice details to UiPath Orchestrator Queue, passing invoice sub totals as an argument  5). Writes the scraped invoice data to an Excel file  6). Calls a workflow to Email an Excel file with the scraped invoice data to an email account | - A list of PDF files has been passed to the Sequence | - Scrapes data from PDF invoices and writes it to an Excel file  - Uploads scraped data to a UiPath Orchestrator Queue  - Sends an email with the scraped invoice data as an attached Excel spreadsheet | - in\_pdfFiles  - out\_InvoiceTable |  |
| Scrape\_PDF\_Sequence.xml | 1). Read a set of PDF files passed to it as an argument  2). Iterates through the PDF's and scrapes the Invoice Date, Invoice Number, Invoice Items, Sub Total, GST %, and Invoice Totals  3). Outputs the scraped invoice item details to Data Tables and the scraped invoice Sub Total details to an Array of Strings | - Receives a set of PDF Files | - Outputs a Data Table of the Invoice Details and an Array of Strings containing the invoice Sub Total Details | - in\_pdfFiles  - out\_FinalInvoiceDetailsDataTable  - out\_SubTotalDetailsArray |  |
| Upload\_Orchestrator\_Excel\_Sequence.xml | 1). Creates a dictionary to store invoice subtotal information  2). Add the Subtotal, GST %, and Total information to the Dictionary  3). Tries to upload this items in the dictionary to an existing UiPath Orchestrator Queue | - An array of the subtotal information from all invoices has been passed to the Sequence | - Uploads subotal information to an existing UiPath Orchestrator Queue as individual transactions | in\_SubTotalDetailsArray |  |

## Robot Type

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| # | Attended | Unattended | Trigger | Comments |
| 1 | Attended |  | Manual | Users must manually Start the ‘Main.xml’ workflow to run this bot |

## Business Exceptions Handling

The Business Process Owner and Business Analysts are expected to document below all the business exceptions identified in the automation process. These can be classified as:

### Known Exceptions

The table below reflects all the business process exceptions encountered during the process evaluation and documentation. These are known exceptions that occurred before. For each of these exceptions, define a corresponding expected action that the robot should complete if it encounters the exception.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| BE # | Exception Name | Step | Parameters | Action to be Taken |
| 1 | “Directory Not Found Exception” | “Read\_PDF\_Write\_Excel\_Sequence.xml” | Filter, Folder | Confirm that the invoice files are stored in a ‘Data’ folder one level beneath the root in a ‘.pdf’ format. |
| 2 | “FileNotFoundException” | “Scrape\_PDF\_Sequence.pdf” | File Name, Range | Confirm that a list of files has been passed to the sequence an argument |
| 3 | “FileNotFoundException” | “Email\_Excel\_Sequence.xml” | Password, Sender Email Address, Recipient Email Address, Port, Server, Email Subject, Email Body, Attachment | Verify that the invoice file exists int the referenced folder |

### Unknown Exceptions

For all other unanticipated or unknown business (process) exceptions, the robot should:

* Display a generic ‘Process Failed with an Unknown Exception’ message.

## System Exceptions Handling

A comprehensive list of all errors, warnings or notifications should be consolidated here with the description and action to be taken, for each, by the robot.

Errors identified in the automation process can be classified as:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| SE # | Exception Name | Step | Parameters | Action to be Taken |
| 1 | “Mail Exception” | “Read\_Email\_Sequence.xml” | Email Address, Username, Port, Server, Mail Folder | Verify Email and Password, confirm port and server are correct for email client, confirm emails are being read from the ‘Inbox’ |
| 2 | “OrchestratorHttpException” | “Upload\_Orchestrator\_Excel\_Sequence.xml” | Queue name, Item Information, ItemInformationCollection | Confirm that the Queue exists in the UiPath Orchestrator account you are writing to. |

For all the other unanticipated or unknown system exceptions, send an email to **<petersoninquiries@gmail.com>** and attach a screenshot of the error message.

# Other Observations

Include below any other relevant observations you consider needed to be documented here.

* The automation assumes that the ‘Main.xml’ file is stored at the Project root, and all other workflows are contained in a ‘Workflows’ folder.
* The automation will create a ‘Data’ folder to store all invoices if it doesn’t already exists
* The automation will create a ‘Excel\_Files folder to store all Excel files if it doesn’t already exists
* The automation is using IMAP for receiving and SMTP for sending mail. Messages are read from “imap.gmail.com” on Port 993 and sent to “smtp.gmail.com” on Port 465.

# Additional sources of process documentation

* This project is stored at the following location on Github: <https://github.com/jrpete/udacity_invoice_scraping>
* The Project contains a README.md folder describing its contents