



Utilização de Metodologias RPAs na Transformação Digital de Processos Empresariais

BRUNO MARCELO RIBEIRO DUARTE

Junho de 2021

Use of RPAs Methodologies for the Digital Transformation of Business Processes

Bruno Duarte

A dissertation submitted in partial fulfillment of
the requirements for the degree of Master of Informatics
Engineering, Specialisation Area of Graphics and
Multimedia Systems

Supervisor: Nuno Bettencourt
Supervisor Enterprise: José Veiga

Porto, June 29, 2021

Abstract

Currently, with the exponential development of technology, there are more and more companies that want to take advantage of it to optimize the processes of company departments, there are more and more platforms and software that need to connect to be able to share data and interact automatically.

It was in this context that the Application Programming Interface (API) emerged, which is an intermediary between two applications and that allows them to be able to "talk" with each other, but the API becomes very limiting since it only applies to a connection between two applications so if it is necessary to link several applications between them the API may become unusable.

However, with a greater evolution of technology and the need to improve business processes, Robotic Process Automation (RPA) has recently emerged, which is a tool that allows you to recreate tasks performed by humans, thus being able to interconnect different applications. With the need to optimize the financial sector of euPago, it is intended to find the best solution for that same.

This work describes the automation process of several processes performed by employees manually using the UiPath tool. It can be seen that these repetitive and sequential processes can be easily automated.

In short, this project aims to find the solution that best fits the customer's requirements and the processes in need of optimization.

Keywords: API, RPA, Optimization

Resumo

Atualmente, com o desenvolvimento exponencial da tecnologia, há cada vez mais empresas que querem tirar partido da mesma para otimizar processos dos seus departamentos, também há cada vez mais plataformas e softwares que precisam de se interligar entre si para assim conseguirem partilhar dados e interagirem automaticamente.

Foi neste contexto que surgiu a Interface de Programação de Aplicações (API), que é um intermediário entre duas aplicações e que lhes permite "falar" uma com a outra, mas a API torna-se muito limitativa uma vez que só permite entre duas aplicações, portanto, se for necessário conectar várias aplicações a API pode-se tornar inutilizável.

Porém, com uma maior evolução da tecnologia e com a necessidade de melhorar os processos de negócio, surgiu recentemente a Automação de Processos Robóticos (RPA), que é uma ferramenta que permite recriar tarefas executadas por humanos, podendo assim interligar diferentes aplicações. Com a necessidade de otimizar o setor financeiro da euPago, pretende-se encontrar a melhor solução para o mesmo.

Neste trabalho é descrito o processo de automatização de de vários processos executados por funcionários de forma manual com recurso á ferramenta UiPath. É possível observar que estes processos repetitivos e sequenciais conseguem ser automatizados de forma fácil.

Em suma, este projeto visa encontrar a solução que melhor se adapta aos requisitos do cliente e aos processos que necessitam de otimização.

Contents

List of Figures	ix
List of Tables	xi
List of Acronyms	xiii
1 Introduction	1
1.1 Context	1
1.2 Problem	1
1.3 Objectives	2
1.4 Methodology	2
1.5 Structure	2
2 State of the art	5
2.1 Automation techniques	5
2.1.1 Application Programming Interface	5
2.1.2 Robotic Process Automation	5
2.1.3 Automation Techniques - Comparison	6
2.2 Robotic Process Automation Tools	6
2.2.1 UiPath	6
2.2.2 Automation Anywhere	7
2.2.3 Blue Prism	7
2.2.4 Microsoft Power Automate	8
2.2.5 Comparison of technologies	8
2.3 Value Analysis	9
2.3.1 New Concept Development Model	9
2.3.2 Quality Function Deployment	11
3 Analysis	13
3.1 Business Analysis	13
3.1.1 Functional Requirements	13
3.1.2 Non-Functional Requirements	17
3.2 Value Analysis	17
3.2.1 Opportunity Identification	17
3.2.2 Opportunity Analysis	17
3.2.3 Idea Generation and Enrichment	18
3.2.4 Idea Selection	18
3.3 Value Proposition	20
3.4 Quality Function Deployment	21
4 Design	23
4.1 Components	23
4.2 Process	24
4.2.1 Interconnected Processes	24
4.2.2 Initial process	25
4.2.3 Bank Receipts Download	25
4.2.4 Bank Receipts Extraction	26
4.2.5 Invoices Download	27

4.2.6	Invoices Extraction	28
4.2.7	Invoices Print	29
4.3	Mockups	30
5	Implementation	33
5.1	Download Process	33
5.2	Extraction Process	44
5.3	Renumbering Process	48
5.4	Print Process	53
6	Evaluation	57
6.1	Metrics	57
6.1.1	Performance	57
6.1.2	Task execution time	57
6.1.3	Error rate	57
6.1.4	Number of files affected	57
6.1.5	Bounce rate	58
6.1.6	Employee satisfaction	58
6.2	Hypotheses	58
6.3	Evaluation methodology	58
6.3.1	Survey	58
6.3.2	Control/Test Group	58
6.4	Data analysis	59
6.5	Survey Results	60
7	Conclusion	63
	Bibliography	65
A	Value Proposition Canvas	67

List of Figures

2.1	New Concept Development [3]	10
2.2	Matrix Symbols	11
3.1	AHP tree for ideia selection	18
3.2	Value Proposition - Canvas	21
3.3	House Of Quality	21
4.1	Component Diagram	23
4.2	Activity diagram	24
4.3	Initial process diagram	25
4.4	Download process - Bank receipts	26
4.5	Extraction process - Bank receipts	27
4.6	Download process - Invoices	28
4.7	Extraction process - Invoices	29
4.8	Print process - Invoices	30
4.9	Mockup alert message	31
4.10	Mockup input message	31
4.11	Mockup log message	32
5.1	Initial Input Process.	33
5.2	Automatic Login Process	34
5.3	Open Browser Process	34
5.4	Navigate to page and accept cookies message.	35
5.5	Read file and obtain login credentials	36
5.6	Certificates message and password insert.	37
5.7	Insert multi-channel code.	38
5.8	Click on the menus (first step).	39
5.9	Click on the menus (second step).	39
5.10	Manual Date Process	40
5.11	Create variables for yesterday's date.	41
5.12	Click in the text release notes from the first table.	41
5.13	Click in all release notes for the second table.	42
5.14	Get second data table and click in all rows.	42
5.15	Save document.	43
5.16	Check next button in data table.	44
5.17	First step of the extraction process.	45
5.18	Start of the second step of the process.	46
5.19	Middle of the second step of the extraction process	47
5.20	Final of the second step of the extraction process.	48
5.21	First step of the renumbering process.	50
5.22	Creation of the file location variables.	51
5.23	Second step of the renumbering process.	52
5.24	Create initial variables.	53
5.25	Join files.	54
5.26	Send file to printer.	54
6.1	Answer to first question.	60
6.2	Answer to second question.	61

6.3	Answer to third question.	61
6.4	Answer to fourth question.	61
A.1	Value Proposition - Canvas (Maximized)	68

List of Tables

2.1	Comparison of technologies	8
3.1	FRQ0001 - Manual process of login data	14
3.2	FRQ0002 - Automatic process of login data	14
3.3	FRQ0003 - Manual date insertion	14
3.4	FRQ0004 - Automatic date insertion	15
3.5	FRQ0005 - Allow to choose bank account	15
3.6	FRQ0006 - Choose the type of document for the download process.	15
3.7	RQ0007 - Save documents to predetermined folders	16
3.8	FRQ0008 - Connection to API to get Client ID	16
3.9	FRQ0009 - Files download	16
3.10	FRQ0010 - Document extraction	16
3.11	FRQ0011 - Document print	17
3.12	NFRQ0001 - Visual and intuitive interface	17
3.13	AHP Scale Comparison	19
3.14	Comparison Matrix	19
3.15	Comparison Matrix with sum	19
3.16	Normalized matrix	19
3.17	Idea for Costs	20
3.18	Idea for Complexity	20
3.19	Idea for Evolution	20
6.1	Number of invoices download in last 3 months.	59
6.2	Time for download fifty invoices.	59
6.3	Number of bank receipt download in last 3 months.	59
6.4	Time for download fifty bank receipts.	59
6.5	Extraction process with RPA solution.	60
6.6	Renumbering process with RPA solution.	60

List of Acronyms

AHP	Analytic Hierarchy Process.
AI	Artificial Intelligence.
API	Application Programming Interface.
FFE	Fuzzy Front End.
IT	Information Technology.
NCD	New Concept Development.
PDF	Portable Document Format.
QFD	Quality Function Deployment.
RPA	Robotic Process Automation.
UML	Unified Modeling Language.
VA	Value Analysis.
VE	Value Engineering.

Chapter 1

Introduction

This chapter presents the project developed during the master's degree in Informatics Engineering, area of specialization in Graphics and Multimedia Systems.

Firstly, it is described the context in which this dissertation fits, then the surrounding problem that gave rise to the realization of the same. Lastly, the objectives and existing methodologies are presented, and next, the structure of this dissertation.

1.1 Context

In the last few years and with the exponential growth of companies, the process optimization factor has been increasingly important to make the most of the work developed in companies, and it is in this context that the term *hyperautomation* appears.

The *hyperautomation* is the extension of process automation to a methodology that combines Artificial Intelligence (AI) tools with Robotic Process Automation (RPA) allowing the ambition to automate almost any task that is performed by a user. But process automation can still go further using bots. Using those, a user can create a completely autonomous process that will perform the tasks that should be performed by humans.

The underlying idea is to increasingly automate the work of service companies to achieve productivity rates similar that what has been achieved, for a long time, in the automation of industrial processes. Bots still manage to reduce the error rate that would exist if this process had to be performed by a human, whilst clearly, they still lose in terms of adaptability and flexibility in handling exceptions. In the future, the term *hyperautomation* will play a relevant role in the world of technologies as it has already been identified by Gartner (a world leader in research and consulting) as one of the 10 strategic Information Technology (IT) trends in 2020 [9]. According to Gartner in 2024, service companies will be able to reduce operating costs between 30%.

The company in which the project will take place is euPago. EuPago is a Portuguese payment institution, accredited and supervised by the Central Bank of Portugal and specialized in providing online payment solutions, thus being the ideal solution for any e-commerce business, from companies to private sales.

EuPago's payment system is quite simple basically, the merchant issues a referral, the customer pays, euPago notifies and transfers the funds to the merchant's account, the institution also features integration with various software such as PHC, Prestashop, Shopify, among others, but also presents a solution via Application Programming Interface (API).

1.2 Problem

The main problem is based on the loss of productivity in companies that use human resources to carry out sequential and repeated processes to carry out the most varied tasks. After this problem, other problems arise that derive from the first, the time loss that these types of processes require from the worker because, for the execution of these sequential and repeated processes, several hours, if not working days, are necessities.

Finally, another aspect that derives directly from the problem described refers to the error rate that these same processes cause and, in addition to this, the recovery rate of these same errors. For all these points, it is crucial to combat this main problem in the daily routine of companies.

This problem happens on a large scale at euPago since it is a payment institution that handles many daily payments and transfers that suffers from the problems mentioned above directly in the company's financial area, both in terms of bank receipts issued for each transfer made and for the invoices generated that need to be worked on by employees.

1.3 Objectives

In this section, the objectives to be reached in this master's thesis are defined.

- Analysis of the problem and state of the art study to solve the existing problem;
- Analysis of the company's workflow to be able to automate the maximum number of processes related to the financial area;
- In-depth analysis of the type of tools to be used (RPA);
- Analysis of the tool to be used (UIPath);
- Development of a mechanism for solving the problem using the tool in question;
- Evaluation of the created solution.

1.4 Methodology

With the objectives in view, the work methodology will be as follows:

1. **Problem analysis** (section 1.2) - where the problem is exposed and where it is initially analyzed;
2. **Technology analysis** (section 2.2) - analysis of technologies that help or may help to solve the problem previously exposed;
3. **Value analysis** (section 3.2) - where a study is made of the value the solution presents;
4. **Design** (chapter 4) - where the design of the solution is built based on customer requirements;
5. **Implementation** (chapter 5) - based on the chosen technology and the designed design, the solution is implemented;
6. **Evaluation** (chapter 6) - where an evaluation plan of the implemented solution is created;
7. **Result analysis** (section 6.4) - comparison between the results obtained with this new solution and the results obtained without it.

1.5 Structure

In this section, the entire structure of this thesis is described.

The first chapter is the introduction, where the project and the company where this project is developed are briefly presented.

In addition to these, the problem that gave rise to the thesis is presented, as well as the objectives that are intended to be achieved with it. To finish this introduction, the following methodologies are presented, as well as the structure of the document.

The second chapter is the state of the art, where the automation techniques are described, then the most used tools of the technique previously chosen for the realization of the solution are presented. Finally, in this chapter, the methods that will be used for the value analysis are presented, as well as a comparison between the tools previously described.

The third chapter is the analysis, in this chapter, the business analysis and the value analysis of the entire project are made.

In the business analysis, the requirements are presented by the client as well as the processes that are executed are described.

In the value analysis, the value of this project is analyzed using the methods presented and described in the previous chapter.

The fourth chapter is the design of the solution, here all diagrams and mockups are made to be able to design the application and understand each of the processes previously described.

The fifth chapter concerns the implementation of the solution designed in chapter 4.

The sixth chapter is the evaluation, which describes the entire process of evaluating this solution to be implemented.

Finally, a conclusion is made where the project conclusions are drawn.

Chapter 2

State of the art

As mentioned in chapter 1.2, the basic problem that surrounds this dissertation is based on the loss of productivity in companies that use human resources to carry out sequential and repeated processes to carry out the most varied tasks.

Not everything can be automated and that is why a question that is no longer new is generated "What should be automated and what should be done by humans?" but with the evolution of RPA and AI technologies, there is always a need to reevaluate this same question and review whether the answer remains the same [6].

The technology has predefined processes for different areas, such as finance or accounting, since these are the areas that present the most problems in terms of repetitive processes and with the highest error rate.

But it also has processes that require more time, such as data scrapping that retired data from different documents since this technology has integration with PDF, Word, Excel, but also with websites where it can do the same with data from all web pages that the user wants to take.

In addition to these processes, RPA also achieves full integration with the operating system where it is running, that is, it allows you to perform tasks that occur on the system itself, such as renaming folders/files, moving folders/files, among others.

The RPA technology also achieves integration with root code, exemplifying, integrating Python code with our bot.

The processes described above can be integrated and work together to create the desired bot in this way.

In this chapter, the technologies that can be used to develop the solution the existing problem is described and compared.

2.1 Automation techniques

When companies enter the field of automation, they find two technologies that, although different, present themselves as valid solution.

These two technologies are RPA and API, but the two are quite different.

2.1.1 Application Programming Interface

An API is a tool that makes it possible through code to interact an application with another one.

The data of an application is also present in the other, thus making them interact.

API are widely used in companies since they allow data to be aggregated between several different applications, thus allowing information sharing.

Simply the API work focus is on the back-end interaction, as is the case of sharing information and data.

2.1.2 Robotic Process Automation

On the other hand, RPA interacts between applications through the front-end. Bots use visual interaction forms such as clicks and writing, just like a human user would do to interact between applications, although much faster and with fewer errors.

The most common to occur are these employees who, when working with this type of repetitive processes based on rules, ended up making mistakes. RPA has many benefits for both companies and employees.

In a study carried out for 16 case studies, it was possible to obtain a financial return that varies between 30 and 200 percent in the first year of companies when using this technology [5].

But, it can offer much more than the financial aspect, as it is an example to carry out the processes faster and cheaper since the company can assign more complex cases that need a deeper analysis to the employees who were freed from the tasks basic due to the bot that performs the basic and often boring tasks that were necessary to do.

The RPA also helps in regulating the company's stress as it removes a lot of repetitive work thus allowing the company to really force itself on important and complicated tasks.

Therefore, the RPA demonstrates benefits for all those involved in the company as well as the customers who will be served more quickly [5].

2.1.3 Automation Techniques - Comparison

Although these two technologies are quite different since one works with the front-end and the other with the back-end, the two complement each other and are often used together to draw the strengths of one and the other.

The question that arises when talking about these two technologies is which one should I use? and this answer always depends on the objective of the project.

Simply, the API allows us to quickly connect two applications.

But, on the other hand, RPA does allow the expansion of the number of applications that can be connected.

Therefore, if the objective is to connect only two applications, the solution must use the API, while if the objective is to connect an undefined number of applications, RPA should be used.

Two of the points that interfere in the choice of the method to be used are the price and the associated cost. Building an API solution is more expensive and takes longer than building a solution using RPA.

But, in terms of maintaining the same changes, in this field the RPA is becoming more expensive since it has to be paid for a software monthly that allows us to build the RPA for corrections, however small they may be, thus leaving us dependent on this software [10].

In conclusion, RPA becomes the technology that presents itself best, so it is the technology used to build the solution to the initial problem.

2.2 Robotic Process Automation Tools

Nowadays, and with the constant evolution of RPA technology, more and more companies are building RPA software. Below, are analyzed the most relevant RPA technologies UiPath, Automation Anywhere, Blue Prism, and Microsoft Power Automate. These four technologies have the same structure, both have three types of bots that are available to the user.

2.2.1 UiPath

UiPath is a software company that develops RPA tools that help companies automate their processes.

It is also relevant to note that UiPath has two editions Enterprise and Community. Enterprise edition is suitable for companies as this version includes all three parts of UiPath. Community edition can be used to learn UiPath since it has a free period and it is also possible to have a paid version of this version for small businesses that don't need all parts of UiPath [7].

This RPA tool consists of three parts:

- **UiPath Studio**, where it is possible to design the entire process you want to automate, it is not necessary to have experience with the code for its design because it presents a visual interface of the process;

- **UiPath Robot**, where it is possible to create bots in order to automate the process designed in UiPath Studio. In UiPath Robot it is possible to choose between attending (works only with human interaction) and unattended (works completely autonomously);
- **UiPath Orchestrator**, where these same bots can be managed and run on the desired machine as well as the entire automation process.

Below are shown the three types of bots that can be created with UiPath.

- **Attended** - This bot operates with the help of a human and it is impossible to start these bots from Orchestrator because these bots are not totally autonomous since they need human intervention such as clicking on certain options;
- **Unattended** - This bot, unlike Attended, works completely autonomously, it can also be executed remotely through Orchestrator since it doesn't need any human intervention;
- **Free** - This bot is very similar to Unattended, but can only be used in a development and test environment, not in production since it can only be configured or used through a specific UiPath account.

2.2.2 Automation Anywhere

Automation Anywhere is a user-friendly RPA tool that aims to automate all processes that are performed by humans.

Below are three software types that interact with the entire bot process from design to maintenance and use.

- **Bot Creator**, where developers are allowed to create bots to automate their computers;
- **Control Room**, an application that acts as a controller and handles the various executions and scripts, credentials, security, clients;
- **Bot Runner**, a machine that allows user to run the bot. After executing this process, users are sent back to the Control Room where they can see the status of the bot.

Below are shown the three types of bots that can be created with Automation Anywhere [9].

- **Task Bot** - In this bot, repetitive tasks based on rules are automated, as is the case with document administration;
- **Meta Bot** - In this bot, automation is performed in building blocks to facilitate scalability with next-generation integration;
- **IQ Bot** - In this bot, there is continuous learning and improvement so that there is automation of processes.

2.2.3 Blue Prism

Blue Prism is an RPA tool aimed at companies since its main objective is to meet all its needs using scalable, configurable automation and centrally managed (control point) [7].

In a simple way, Blue Prism consists of three components:

- **Application Server**, works as the brain of the entire configuration of the Blue Prism process through the execution of functions such as connecting to databases and storage of logs [8];
- **Runtime Resource**, or better known as bots could be both physical and virtual computers [8];
- **Interactive Client**, is a tool installed on your own computer that allows humans to be able to interact with their bots through the development of processes and tasks [8].

Blue Prism addresses some types of processes like rules-based processing with digital structured inputs, high transaction volumes and others [4].

2.2.4 Microsoft Power Automate

Power Automate is used to automate common business processes or sequences based on several conditions or scenarios. It is a web-based tool designed to interact with a growing library of software from Microsoft and other vendors.

Microsoft Power Automate connects to a large number of applications, but if it doesn't connect to a particular application that you want, it is possible to develop our own connection to work with custom applications. You can even use HTTP to interact with any REST-based interface for any application.

Power Automate presents three possible forms of use:

- **End User Web Portal**, where users can create, design, import, save, export, and execute flows;
- **Mobile App**, same as the End User Web Portal but with the difference of being specialized for mobile devices;
- **Admin**, where you can manage all the administration of the Power Automate environment including viewing the number of bot executions as well as data gateway configurations performed.

Microsoft Power Automate is a cloud-only solution, it does not exist in a computer application version [11].

2.2.5 Comparison of technologies

In this section, a comparison is made between the technologies through a summary table so that it is possible to analyze the best solution.

TABLE 2.1: Comparison of technologies

Requirements	UiPath	Anywhere	Blue Prism	Power Automate
Annual Price Very Low (< 3,000€); Low (5,000€ < > 3,000€); Moderate (10,000€ < > 5,000€); High (15,000€ < > 10,000€); Very High (> 15,000€);	Very High	Moderate	High	Moderate
Free trial	Yes	Yes	Yes	Yes
Free Version	Yes	No	Yes	No
Code-free Development	Yes	Yes	No	Yes
Automated Testing	Yes	No	No	No
Third Party Integration	Yes	No	No	Yes
Number of Supported Platforms	4	4	1	5
Support	Yes	Yes	Yes	Yes
Data Entry	Yes	No	No	No
Unattended Automation	Yes	Yes	Yes	No
Attended Automation	Yes	Yes	Yes	No
Data Extraction	Yes	No	No	No
Image Recognition	Yes	Yes	Yes	No
OCR	Yes	Yes	Yes	No
Workflow Management	Yes	Yes	No	No
Workload Automation	No	No	No	Yes
Cloud-only solution	No	No	No	Yes

After all the technologies are compared, we quickly realized that UiPath is the tool that can be the best alternative for a human user, since it has some specific requirements, such as data entry and extraction, for example, these functionalities are required in RQ02, RQ03 and RQ12.

2.3 Value Analysis

Value Analysis (VA) is a crucial phase in the process of creating a new product.

VA can be defined as a systematic review process applied to the product to be developed or to the product already developed. To compare the product function required by a customer to meet its requirements at the lowest cost consistent with the specified performance and required reliability [1] .

In order to simplify the definition of value analysis, it can be divided into five key points:

1. VA (and Value Engineering (VE)) is a systematic, formal, and organized process analysis and evaluation, which requires planning, control, and coordination;
2. The analysis concerns the function of a product to meet the demands or application necessary for a customer;
3. Understand the use of a product to assess the level of adequacy between that product and the value obtained by the customer or consumer;
4. To have success, the formal management process must meet these functional specifications and performance criteria consistently;
5. The review process should result in a process of design improvements to reduce production costs while maintaining the level of value to generate benefits for the company.

The VA approach, as process, comprises five stages:

1. Orientation;
2. Functional Identification/Analysis;
3. Creation of Alternatives;
4. Analysis and Evolution;
5. Implementation.

2.3.1 New Concept Development Model

New Concept Development (NCD) Model provides a common language and definition of key components associated with the Fuzzy Front End (FFE).

This concept is divide into three fundamental areas.

- Engine - consists of the core elements that provide power to the front-end process.
- Wheel - consists in five elements keys (Opportunity Identification, Opportunity Analysis, Idea Generation, Idea Selection, Concept Definition).
- Rim - consists in the external environmental factors that influence the engine and the elements.

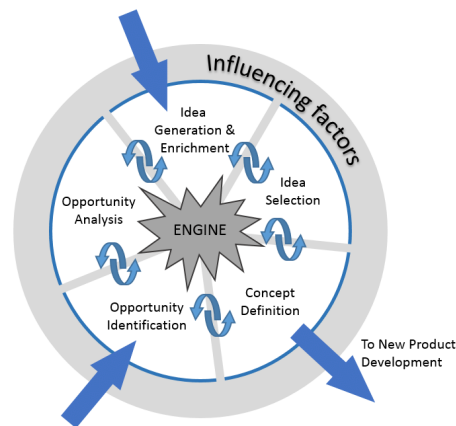


FIGURE 2.1: New Concept Development [3]

To be able to build a new product/solution, it is necessary to define the five elements present in the wheel.

- Opportunity Identification - This element aims to identify exploitable opportunities. To perform this there are different techniques that enable its execution [3]:
 - Roadmapping - capture the strengths of a business to enable mapping in order to improve knowledge sharing.
 - Technology Trend Analysis - research and analysis of technological trends that offer competitiveness to the product.
 - Customer Trend Analysis - research and analysis of customer trends.
 - Competitive Intelligence Analysis - aims to obtain and analyze information about competitive trends outside the company, thus making it possible to develop a business strategy.
 - Market Research - potential customers who need an RPA tool that will help reduce the time lost by employees are taken into account.
 - Scenario Planning - scenarios are developed to get an idea of the future of the project.
- Opportunity Analysis - The first element to be evaluated by companies is the opportunity analysis because it is through this that an organization can identify the opportunities that the company wants to pursue [3].
- Idea Generation and Enrichment - Once the opportunity and validity are defined, the next phase involves presenting the ideas.
- Idea Selection - After create the initial ideas, all the ideas defined will now be evaluated to reach the selection of the best possible idea.

To help this selection of ideas, the Analytic Hierarchy Process (AHP) method will be used. AHP is a method that allows you to make complex decisions based on significant attributes and considering the different alternatives.

- Concept Definition - The concept definition is the final element of the NCD model, which is the only one that presents an exit to this model. After the approval of the previously selected idea, it is possible to create a process that will lead to the success of this solution.
 - The financial sector workflow will be analyzed to automate the maximum number of processes;
 - A solution will be created that meets all previously established requirements;
 - Evaluation of the solution created to compare the initial requirements with the functions present in the solution that fulfill them.

2.3.2 Quality Function Deployment

Quality Function Deployment (QFD) is a technique designed to ensure that customer needs are always taken into account throughout the project.

The QFD involves four matrices:

- House of Quality - convert customer attributes into engineering characteristics.
- Parts Deployment - convert engineering characteristics into parts characteristics.
- Process Planning - convert parts characteristics into process operations.
- Production Planning - convert process operations into production requirements.

The construction of the House of Quality involves the following steps:

- On the left side of the House of Quality, all important customer requirements will be placed.
- Horizontally above the relationship matrix, product design requirements will be added.
- In the relationship matrix, it will be identified which requirements of the design affects the requirements of the customer.
- The correlation matrix will determine how design requirements help and hinder each other.
- Finally, the competitiveness assessment shows how each customer's requirement is met by the competition.

As previously said, some symbols are used in the relationship matrix that shows how strongly each of the design parameters affects the customer's requirements, below, are shown the same symbols.




Symbol	Value	Definition
	9	Strong
	3	Medium
	1	Weak
	0	No assignment

FIGURE 2.2: Matrix Symbols

On the other hand, a correlation matrix will be filled with the symbols "+" and "-" depending on whether the design requirements help or hinder each other.

Chapter 3

Analysis

This chapter is divided into two sections, the first refers to business analysis, and the second to value analysis. This division is carried out as it makes it easier to understand the solution to be implemented.

3.1 Business Analysis

In this section, the processes that are intended to be created are presented.

The objective is to communicate/interact with a banking website to obtain the files that are generated (bank receipts and invoices) and to replace human repetitive and time-consuming processes.

There are two types of files, bank receipts and invoices, each of which will have different processes.

The invoice is a document that must be issued whenever a good or service subject to VAT is purchased, even if it is not requested by the customer, while a bank receipt is a document that contains a summary of the transaction details.

Below are presented all the processes.

- Bank receipts:
 - Download process - this process allows the download of all bank receipts grossly since each file will contain several bank receipts to carry out this process as quickly as possible;
 - Extraction process - this process allows the extraction of all bank receipts and ordering them by adding each of the receipts to selected folders with the id of the customer to which they belong.
- Invoices
 - Download process - this process allows the download of all invoices.
 - Extraction process - this process allows the extraction of the invoice-receipt that are mixed with the invoices. The purpose of this process is to extract documents with the words 'Invoice-Receipt' as these documents serve as receipts for customers.
 - Print process - this process allows the printing of all receipt invoices previously extracted. This process creates a folder where all receipt invoices that have not yet been printed are saved and automatically sends all these receipt invoices to the printer.

In addition to these processes related to each of the files, there is also the initial process where the browser will be opened and the manual or automatic login process on the banking website.

These documents are downloaded both for the company's use, to keep the documents recorded, as well as for customers who want proof of their payments.

3.1.1 Functional Requirements

In this section the customer's functional requirements are presented.

TABLE 3.1: FRQ0001 - Manual process of login data

REQ0001	Manual process of login data
Priority	Essential
Description	It is necessary that the software allows the input of data such as username, password, multichannel code in the bank webpage, these values are entered by a human.
Motivation	This being a software that will have a bot as an intermediary, is necessary that it allows several methods of data insertion in the bank webpage and in this case of the manual process, the user can use the inputs as if were writing himself in the webpage.
Additional Information	The bot will have data entry boxes to enter the username, password and multichannel code.

TABLE 3.2: FRQ0002 - Automatic process of login data

REQ0002	Automatic process of login data
Priority	Essential
Description	The bot will support the insertion of automatic login data on the bank's webpage, the bot will read a config.ini file where the user's data will be and will automatically insert it in the right fields as if it were the user.
Motivation	The main reason for entering this requirement is the time it would take the user to enter all the data, so the process will become faster.
Additional Information	A button will be displayed that will allow the user to automatically insert the data if of course there is the same data in the config.ini file at the root of the project.

TABLE 3.3: FRQ0003 - Manual date insertion

REQ0003	Manual date insertion
Priority	Essential
Description	The bot will support manual date entry. The bot will show inputs where the user should write the range of dates where wants to download the documents.
Motivation	The main reason for insertion this requirement is the fact that the webpage needs to know the date period for which the user wants to download the desired documents.
Additional Information	A data entry box will be displayed and will allow the user to enter the desired period.

TABLE 3.4: FRQ0004 - Automatic date insertion

REQ0004	Automatic date insertion
Priority	Essential
Description	Automatic date insertion will allow the user to save time when entering the date. This date will be inserted by the bot automatically and will put the date of the day before the day that the user is downloading the files.
Motivation	The main motivation for the insertion of this requirement in the system is saves the user time.
Additional Information	The bot will insert the day before the day of download of the files because the banks always work with records from the previous day and not with the current day.

TABLE 3.5: FRQ0005 - Allow to choose bank account

REQ0005	Allow to choose bank account
Priority	Essential
Description	The bot should allow the user to choose the desired bank for the process.
Motivation	This way the user does not need to change the process whenever wants to change the bank to carry out the process.
Additional Information	A message box with a radio button will be displayed for the user to choose the desire bank to act upon..

TABLE 3.6: FRQ0006 - Choose the type of document for the download process.

REQ0006	Allow to choose the type of document for which want to run the process
Priority	Essential
Description	The bot should allow the user to choose the desired type of document for the process.
Motivation	This way the user does not need to change the process whenever he wants to change the type of document to carry out the process.
Additional Information	A message box with a radio button will be displayed for the user to choose the desire type of document.

TABLE 3.7: RQ0007 - Save documents to predetermined folders

REQ0008	Put documents in predetermined folders
Priority	Essential
Description	The bot should save the transferred documents inside predetermined folders in order to facilitate the visualization of each document.
Motivation	The reason for the insertion of this requirement is based on the ease of consulting each of the documents as well as whom is associated with it.
Additional Information	Not Applicable.

TABLE 3.8: FRQ0008 - Connection to API to get Client ID

REQ0009	Connection to API to get Client ID
Priority	Essential
Description	The bot must connect to a base API that will receive the data sent by the bot and return the client ID via an HTTP request to write the client ID as the default folder name.
Motivation	That insertion of this requirement is since it becomes more noticeable and easier to read the documents if we know who refers to the folder where the documents will be found.
Additional Information	This process will take place in the background without the user's perception.

TABLE 3.9: FRQ0009 - Files download

REQ00010	Documents download
Priority	Essential
Description	The bot should download several files consecutively.
Motivation	Shorten the download process to the file destination folders as much as possible.
Additional Information	Not Applicable.

TABLE 3.10: FRQ0010 - Document extraction

REQ00011	Documents extraction
Priority	Essential
Description	The bot should allow extraction from downloaded files.
Motivation	Make it easier for employees who have to review documents.
Additional Information	This process should work in the background.

TABLE 3.11: FRQ0011 - Document print

REQ00012	Documents print
Priority	Essential
Description	The bot must allow the paid and extracted invoices to be printed.
Motivation	Make sure that employees do not waste a lot of time sending documents to the printer.
Additional Information	The bot should gather all documents that have not yet been downloaded and send them to the printer.

3.1.2 Non-Functional Requirements

In this section the customer's non-functional requirements are presented.

TABLE 3.12: NFRQ0001 - Visual and intuitive interface

REQ0007	Visual and intuitive interface
Priority	Essential
Description	The bot should have an intuitive interface and feedback messages on the process status.
Motivation	The reason for the insertion of this requirement is since the user needs to know where the process of extraction or download documents is.
Additional Information	Not Applicable.

3.2 Value Analysis

In this section, all the methods described in 2.3 are used in order to obtain an analysis for the value of the solution.

3.2.1 Opportunity Identification

Regarding the solution to be implemented, it was not found in the market, this aspect occurs because this solution has customer-specific requirements.

In the problem presented in section 1.2, the need to reduce the time spent on repetitive tasks by employees is identified. To this problem, it is possible to add yet another one referring to the errors that these tedious processes cause in carrying out the process.

As mentioned in the section 2.2, there are several RPA technologies that allow the development of this application.

Regarding the technology used, UiPath will be used because it is the one that best fits this solution.

3.2.2 Opportunity Analysis

This is a project aimed for a company, so this solution presents itself as an opportunity for the company to be able to optimize the work of its employees, as well as reduce the errors that they may make and not forgetting the aspect of removing the most boring tasks from day-to-day work.

From the point of view of analyzing opportunities for other companies, this project is also very important, as it was not possible to find similar solutions.

Even if the solution cannot fully fit the ideas of these companies, it will always be possible to adapt the solution to them and this is one of the advantages of using RPA tools.

3.2.3 Idea Generation and Enrichment

Faced with the problem, two initial ideas were created by the company.

The first idea is to use an API that connects different applications through a back-end, while the second idea is to use an RPA technology that connects applications through their front-end.

These two ideas aim to execute all the processes created in the business analysis.

The next step is to evaluate each of these generated solutions and thus reach a selection of one of the ideas.

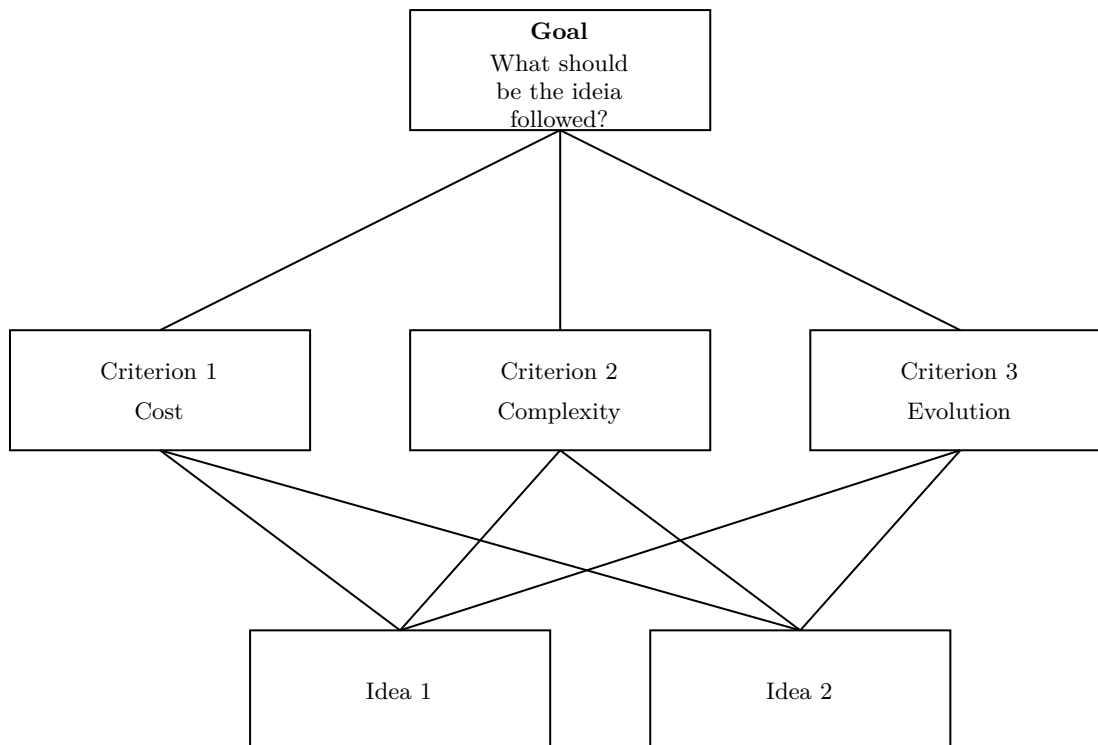
3.2.4 Idea Selection

The first step in building an AHP model to arrive at an idea selection is to design an AHP tree, for this, we first need to define the attributes we will use to evaluate all possible alternatives for creating the solution, namely:

- Cost - This attribute defines the price factor for each of the alternatives;
- Complexity - This attribute defines the complexity that each of the alternatives presents;
- Evolution - This attribute defines how easy it will be in the future to evolve this solution using the software used;

Now that the attributes to be used are defined, the next step is to design the AHP tree.

FIGURE 3.1: AHP tree for idea selection



After designing the AHP tree for the selection of ideas, the next step is to create a scale in table 3.13 to be able to evaluate each of the initial ideas.

TABLE 3.13: AHP Scale Comparison

Scale	Preference Degree
1	Equal Importance
3	Moderate importance
5	Strong or essential importance
7	Very strong importance
9	Extreme importance
2,4,6,8	Values for inverse comparison

Once the scale used to evaluate the attributes has been demonstrated, the respective comparison matrix is created in 3.14 and in table 3.15.

TABLE 3.14: Comparison Matrix

	Costs	Complexity	Evolution
Costs	1	1/3	5
Complexity	3	1	3
Evolution	1/5	1/3	1

TABLE 3.15: Comparison Matrix with sum

	Costs	Complexity	Evolution
Costs	1	1/3	5
Complexity	3	1	3
Evolution	1/5	1/3	1
Sum	21/5	5/3	9

Once the comparison matrix has been defined and presented with the sum of the criteria, the next step is to make the matrix normalized in table 3.16.

TABLE 3.16: Normalized matrix

	Costs	Complexity	Evolution	Priority Vector
Costs	0,2381	0,2000	0,5556	0,3312 (33%)
Complexity	0,7143	0,6000	0,3333	0,5492 (54%)
Evolution	0,0476	0,2000	0,1111	0,1195 (12%)
Sum	1	1	1	

The next step is to assess the consistency of the previous criteria.

$$\lambda_{max} = \begin{bmatrix} 0,2381 & 0,2000 & 0,5556 \\ 0,7143 & 0,6000 & 0,3333 \\ 0,0476 & 0,2000 & 0,1111 \end{bmatrix} \cdot \begin{bmatrix} 0,3312 \\ 0,5492 \\ 0,1195 \end{bmatrix} = \begin{bmatrix} 3,3579 \\ 3,4624 \\ 3,0849 \end{bmatrix} \quad (3.1)$$

$$\lambda_{max} = ((3,3579/0,3312) + (3,4624/0,5492) + (3,0849/0,1195))/3 = 3,3017 \quad (3.2)$$

$$IC = (3,3017 - 3)/(4 - 1) = 0,1509 \quad (3.3)$$

$$RC = (0,13)/(0,9) = 0,1676 = 16,76\% \quad (3.4)$$

In equations 3.1 and 3.2, the largest eigenvalue of the comparison matrix is calculated, which in turn will be used to calculate the values of the consistency index (equation 3.3) and the consistency ratio (equation 3.4).

Now, we have to make the comparison matrix for each alternative with the criteria.

TABLE 3.17: Idea for Costs

	Idea 1	Idea 2	Priority Vector
Idea 1	1/3	1	0,25
Idea 2	1	3	0,75

TABLE 3.18: Idea for Complexity

	Idea 1	Idea 2	Priority Vector
Idea 1	1/3	1	0,25
Idea 2	1	3	0,75

TABLE 3.19: Idea for Evolution

	Idea 1	Idea 2	Priority Vector
Idea 1	1/5	1	0,1666
Idea 2	1	5	0,8333

After the matrices are made, the final equation is presented.

$$\begin{bmatrix} 0,25 & 0,25 & 0,1666 \\ 0,75 & 0,75 & 0,8333 \end{bmatrix} \cdot \begin{bmatrix} 0,3312 \\ 0,5492 \\ 0,1195 \end{bmatrix} = \begin{bmatrix} 0,2400 \\ 0,7599 \end{bmatrix} \quad (3.5)$$

In equation 3.5, the compound priority for the alternatives is obtained, in which the priority matrix is multiplied by the weights of the criteria. Thus, the highest value returned from this multiplication is the idea that presents itself as the most suitable choice.

Once the calculations are finished, we find that the best idea is Idea 2 through the use of RPA technology.

3.3 Value Proposition

The value proposition is an overall view of a company's bundle of products and services that are valuable to the customer [2].

This section presents four questions that must be answered:

- What is your product?
The product is an RPA solution to solve problems such as repetitive and sequential tasks, the same adapting to the needs of the client company.
- Who is your target customer? For whom you provide value?
The target customer is all companies that contain financial sector.
- What value you provide?
The value provided refers to the optimization of sequential and repetitive processes by the employees but also the reduction of errors associated with these same processes. On the other hand, this solution will also provide employees of these companies with the ability to focus on tasks that require more difficulty.
- Why your product is unique?
This solution is unique since it isn't on the market due to the specific requirements it contains.

In figure 3.2 it is presented the VA Canvas for this project.

Analyzing the figure 3.3, we were able to understand the client's requirements as well as their importance in this project. In addition to these aspects, the requirements that will have the most impact in terms of design, as well as the correlation between technical requirements, are also visible.

In this chapter, the business analysis and the analysis of the value of the project were carried out, the system requirements were also defined and the idea that will be implemented and that will be drawn in the next chapter was chosen.

Chapter 4

Design

This chapter aims to detail the design for the solution. This creation and design of the solution will always be based on all the technical requirements presented in chapter 3.

First, a section referring to the components that are interconnected in this solution is presented. Then, processes are then drawn and presented in diagram format. Finally, the mockups of the feedback messages are demonstrated.

All of the following diagrams use the Unified Modeling Language (UML) methodology and are through activity diagrams.

4.1 Components

After analyzing the requirements in the section 3.1, a component diagram is now presented to understand the components involved in the project.

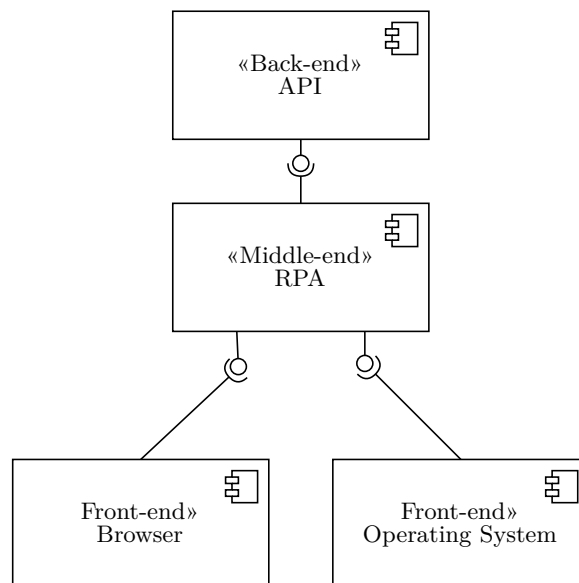


FIGURE 4.1: Component Diagram

In the figure 4.1, it is possible to note that there is an API that will support RP. This API is an API of the company that will provide necessary information about customers to achieve more accurate processes.

In addition to these two components, we can also note that there are two different front-ends, the browser, and the operating system, this division was made since the two have different ways of RPA interacting with the machine.

For the RPA to be able to interact with the browser, it will perform a click/input process in the browser using the extension provided by UiPath for the different browsers.

In the case of the operating system, it is a background process since the solution will be developed on a Windows environment and UiPath has full integration with the same.

Once the overview of the components that make up this project is done, all the existing processes in the system are now described and their diagrams are presented.

4.2 Process

Based on the section 3.2.3 and once the idea of using the RPA tool to build the solution was selected, the processes are now described.

4.2.1 Interconnected Processes

In figure 4.2 shows a diagram of the entire process. This diagram shows all the choices that the user can make as well as all the RPA responses to a given user choice.

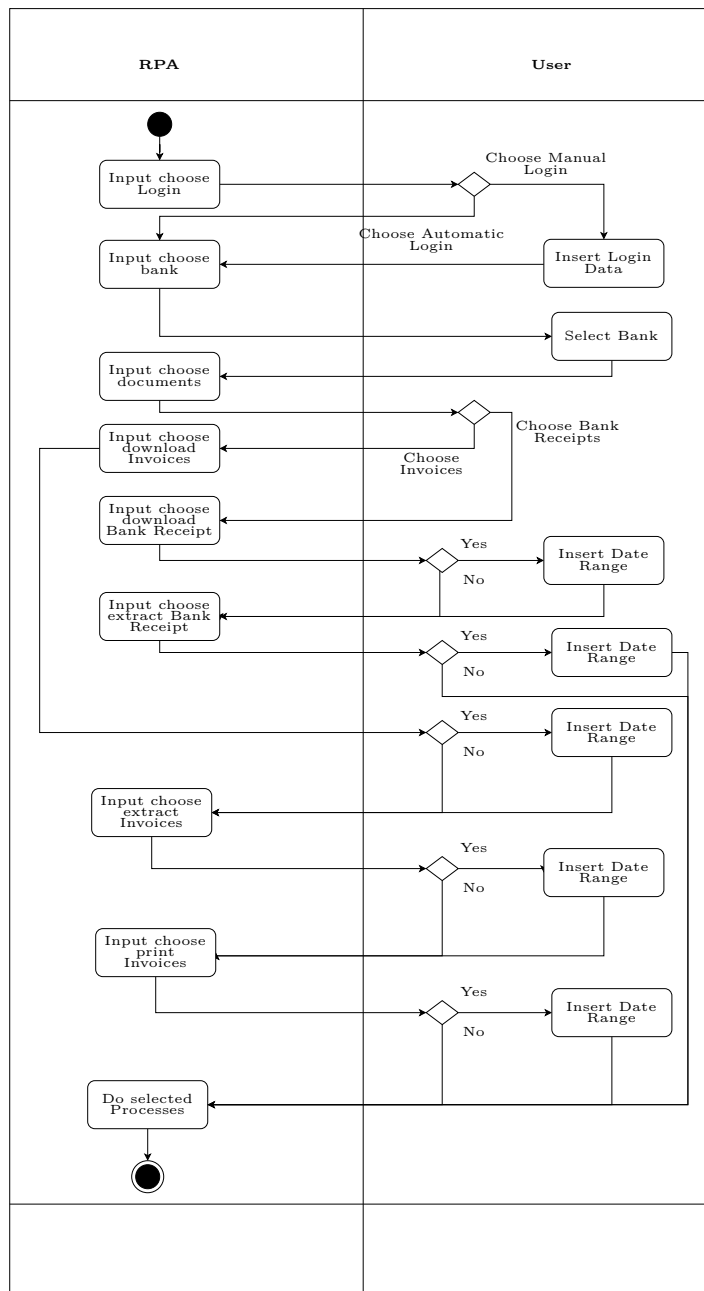


FIGURE 4.2: Activity diagram

In the next sections, each of the processes presented in this diagram will be demonstrated in more detail.

4.2.2 Initial process

The initial process, is a process that opens the browser in which we intend to open the banking website as well as to perform manual or automatic login.

The figure 4.3 is a process diagram showing all the actions in this process.

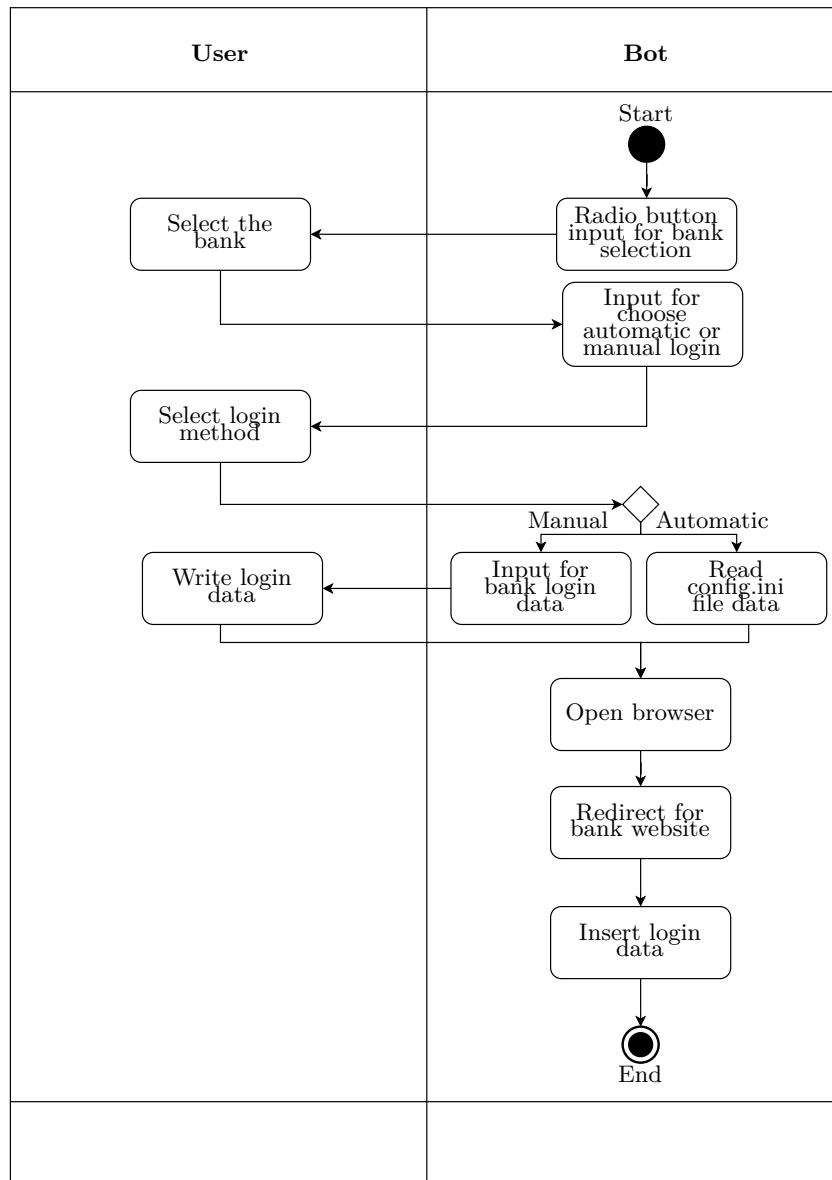


FIGURE 4.3: Initial process diagram

4.2.3 Bank Receipts Download

This process of downloading bank receipts aims to download all receipts in a clustered form, that is, each file will have numerous receipts since the number of this type of documents is quite high and to optimize this time, the download should be done this way.

The download process is a process common to both bank receipts and invoices, so it is very similar.

This download process will be performed for a Google Drive mass mapped.

The figure 4.4 is a process diagram showing all the actions in this download bank receipt process.

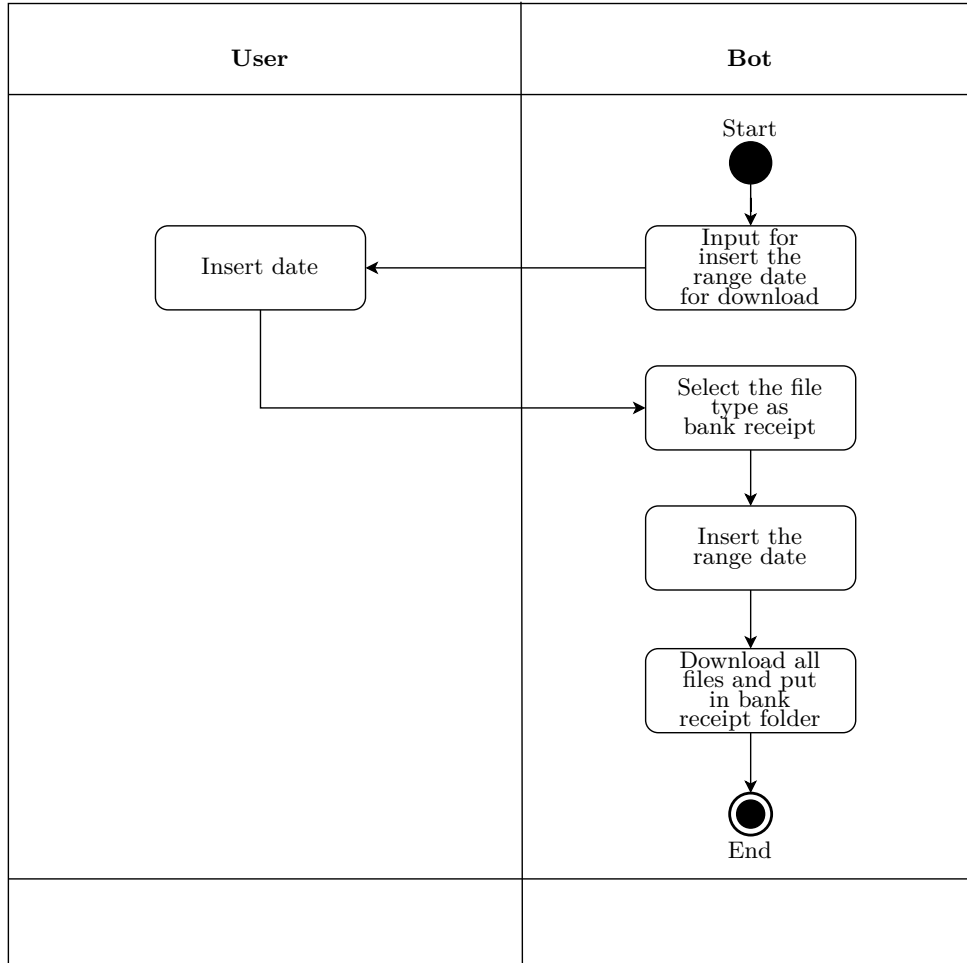


FIGURE 4.4: Download process - Bank receipts

4.2.4 Bank Receipts Extraction

The process for extracting bank receipts aims to extract all documents page by page and create a folder with the client's id where the extracted documents will be saved according to the client to which they relate.

The figure 4.5 is a process diagram showing all the actions in this extraction bank receipt process.

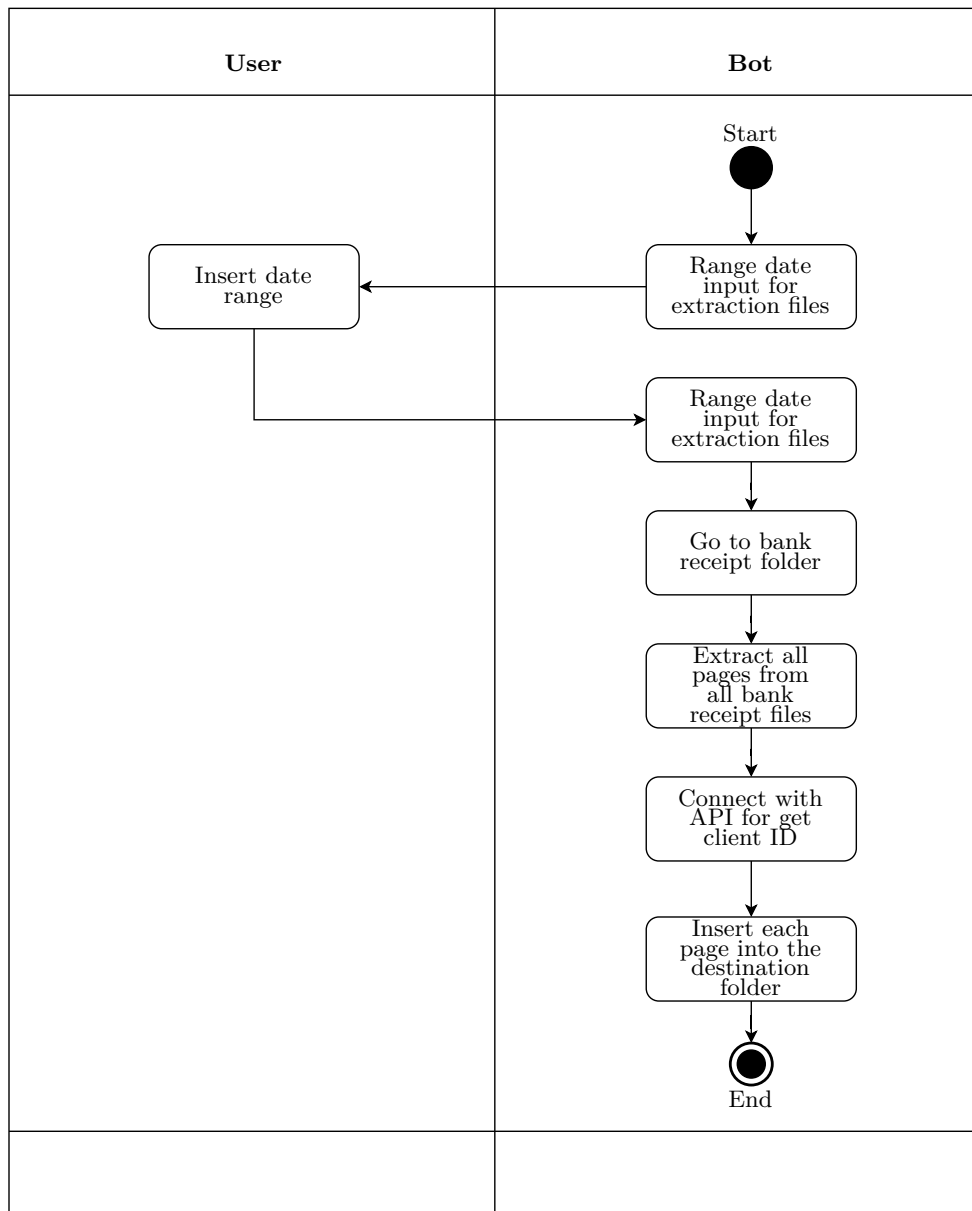


FIGURE 4.5: Extraction process - Bank receipts

4.2.5 Invoices Download

The process for downloading invoices aims, like the process of downloading bank receipts, to download all files but this time for the invoice.

The figure 4.6 is a process diagram showing all the actions in this download invoices process.

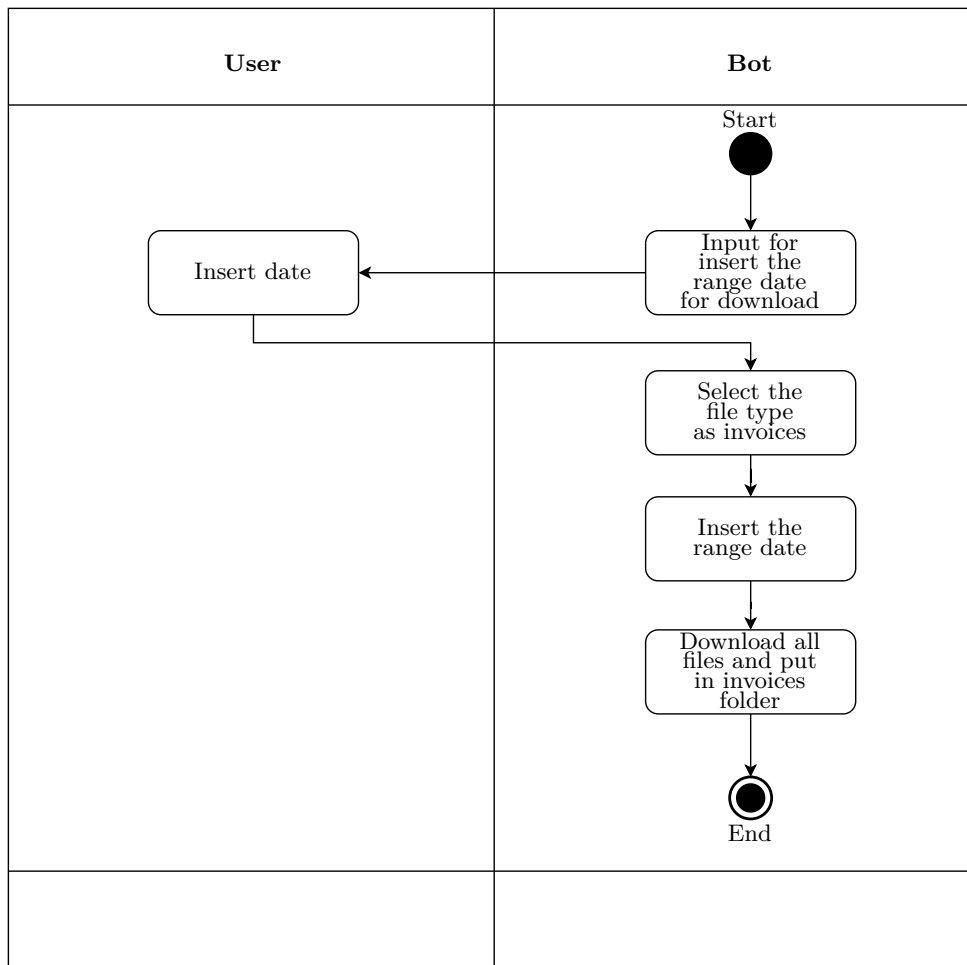


FIGURE 4.6: Download process - Invoices

4.2.6 Invoices Extraction

The process for invoice extraction aims to extract all pages that contain the word "Invoice-receipts" from all files in the invoices folder. In addition to this, the process should also copy these files to the print folder.

The figure 4.7 is a process diagram showing all the actions in this extraction invoices process.

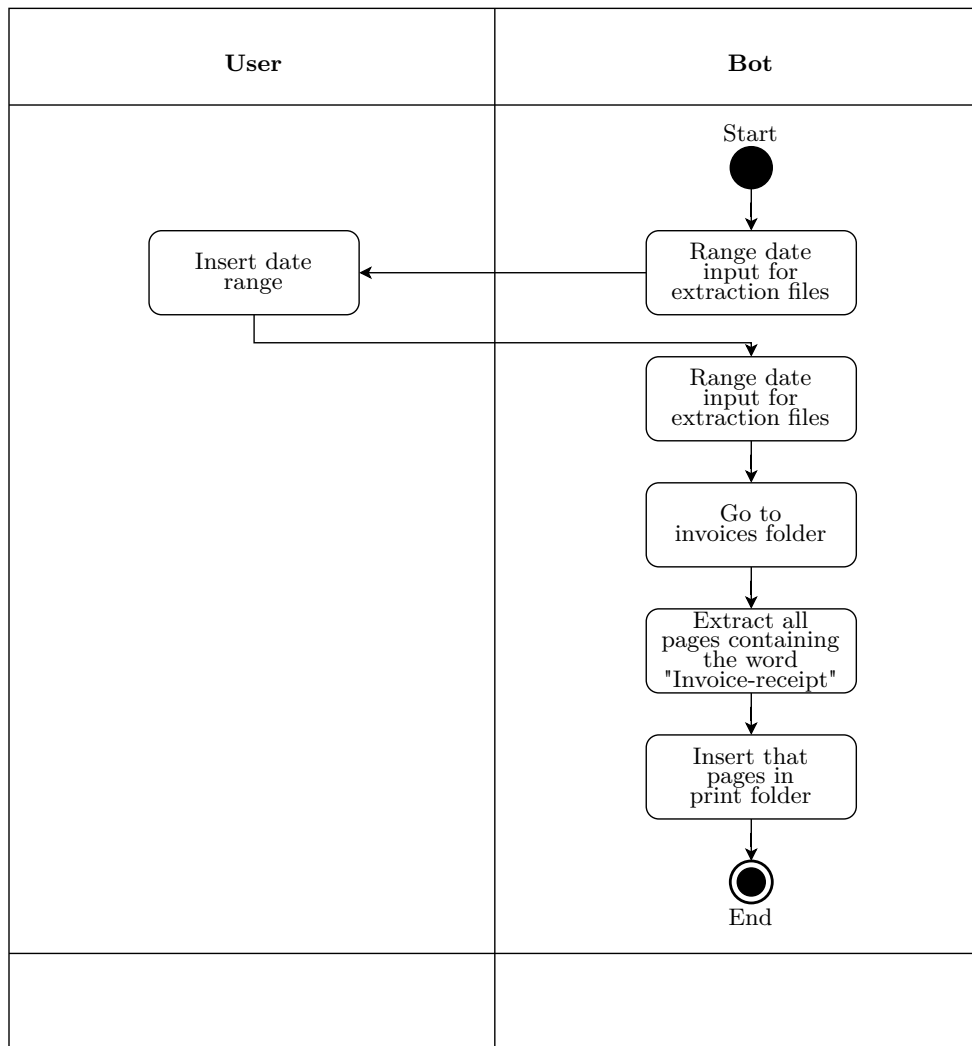


FIGURE 4.7: Extraction process - Invoices

4.2.7 Invoices Print

The invoice printing process aims at printing all files in the print folder and delete them as they are printed.

The figure 4.8 is a process diagram showing all the actions in this print invoices process.

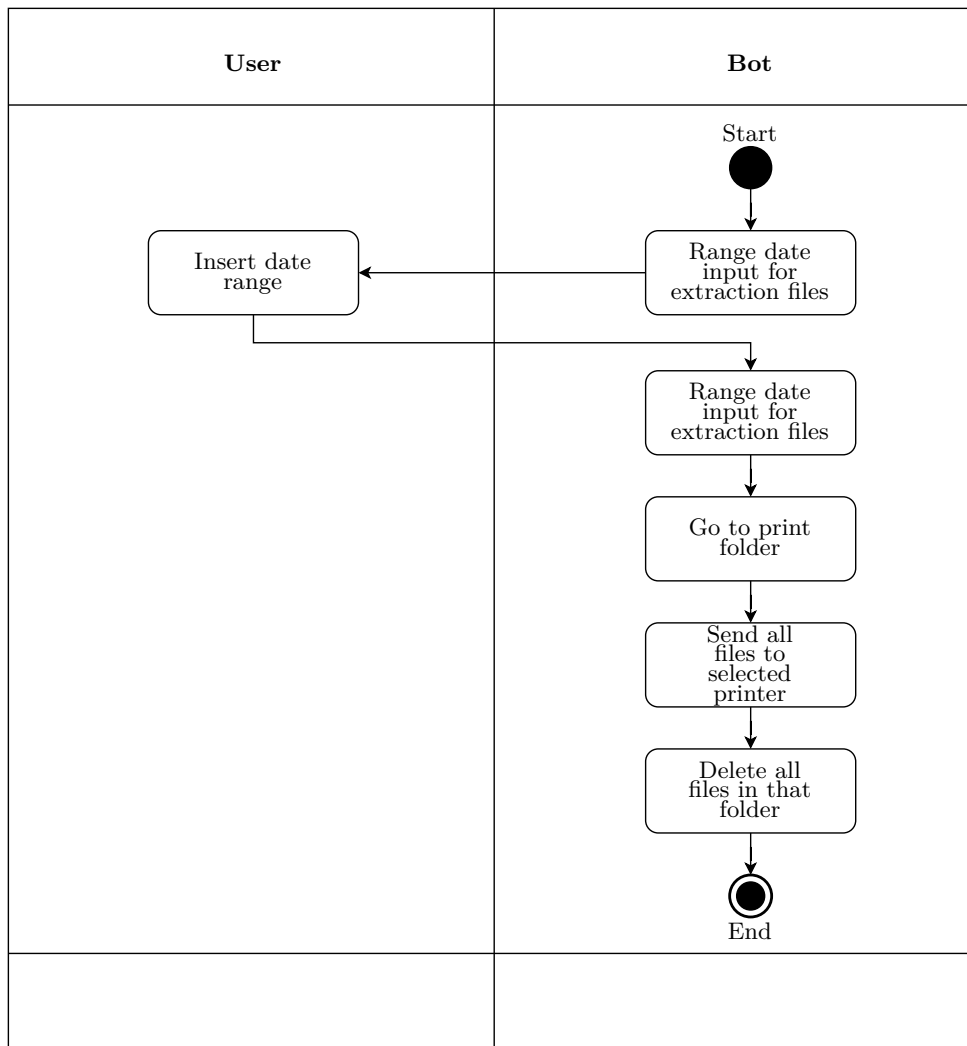


FIGURE 4.8: Print process - Invoices

4.3 Mockups

In this section, the graphical interface that is presented to the user is designed.

From this graphical interface, since the process is executed by a bot that practically runs the processes in the background, the user is shown various types of messages about each of the processes.

Among the previous messages mentioned are input messages for the user to enter data figure 4.10, log messages that appear in the UiPath interface itself figure 4.11, and alert messages when the process ends or when an error occurs figure 4.9.

Below, are shown through mockups each type of message received by the user throughout the process.

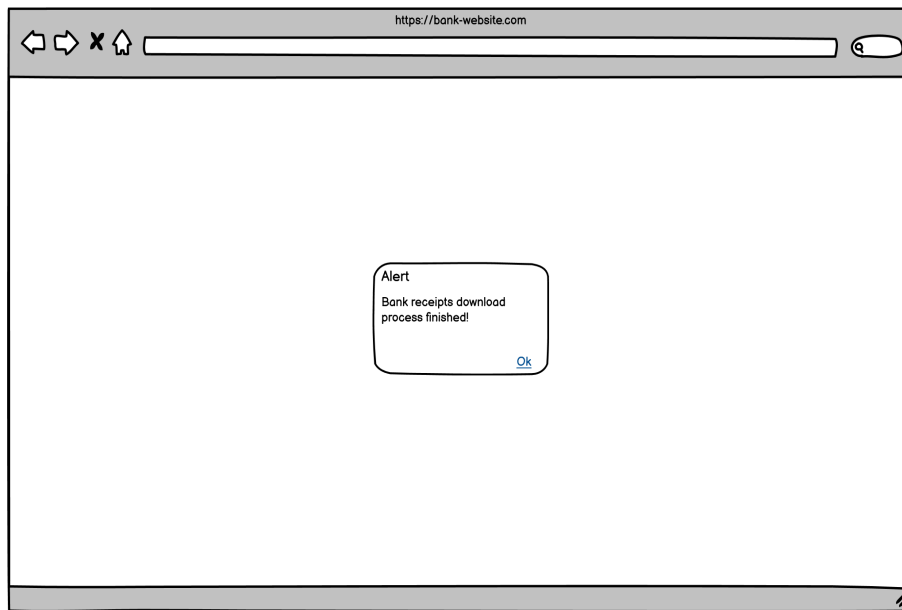


FIGURE 4.9: Mockup alert message

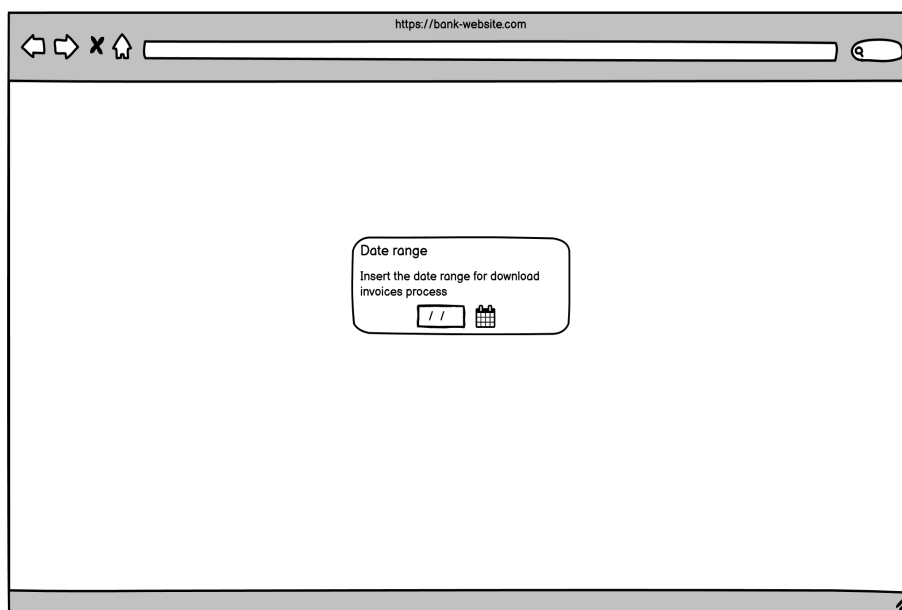


FIGURE 4.10: Mockup input message

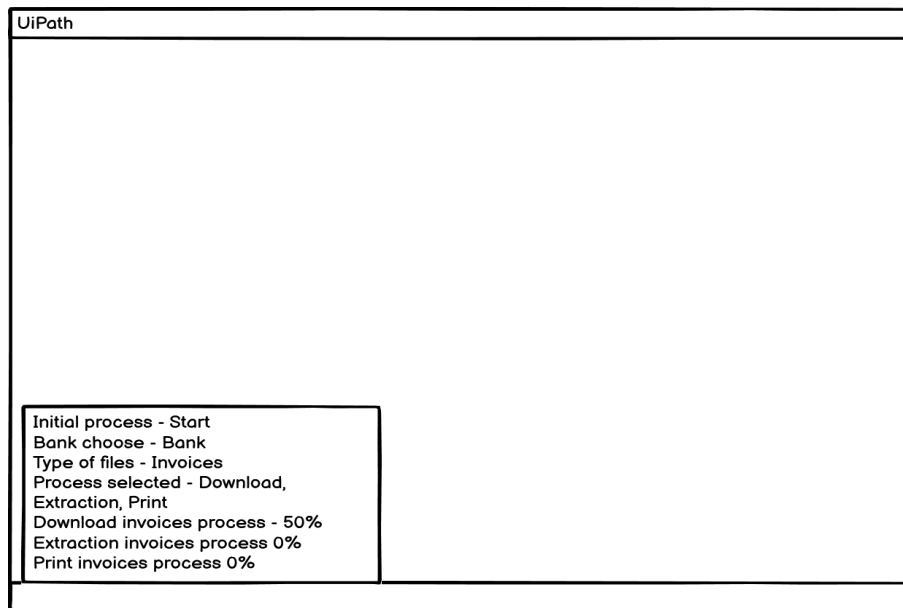


FIGURE 4.11: Mockup log message

Once the entire design of the solution has been presented, from the design of the processes, the design of the feedback messages, and the components that involve this solution, the implementation of this design is presented in the next chapter.

Chapter 5

Implementation

This chapter demonstrates the entire implementation of the main processes that exist in the solution.

This chapter is divided into sections that refer to each of the processes presented in 4.2. These diagrams are presented in UiPath application screenshot formats.

5.1 Download Process

In this section, the implementation of the download process is carried out.

As foreseen and designed in subsection 4.2.5, the purpose of this process is to download all existing invoices for the intended date so that other processes can be applied to them.

Based on the diagram 4.6, it is possible to divide this process in three steps.

The first step consists of the data entries entered by the user regarding the dates as well as the login data for the bot to be able to log into the banking platform.

As for the second step, it is constituted by the process of login process. This process involves logging into the banking platform using the data entered by the user or using the predefined data for login. These same data corresponds to the username, password, and multi-channel code.

In the third step of this process the download process. This process consists of downloading all documents on the dates previously chosen by the user.

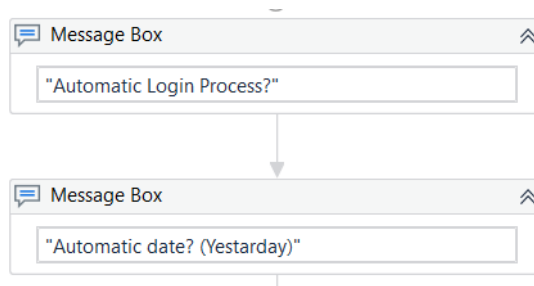


FIGURE 5.1: Initial Input Process.

In the initial process, the user has to choose whether wants automatic or manual login and the same applies to the date, figure 5.1.

In automatic login, the bot uses the predefined data in a "config.ini" file located in the project's root folder. In this file, you can find the necessary login data such as username, password, and multi-channel code.

As far as the automatic date process is concerned, it refers to the day before the current date the bot is running, that is, the user will not have to enter any date as the bot will work with the date of the day previous automatically.

In the manual login, the user will have to enter the data manually in inputs that will appear.

In the manual date, as in the manual login, inputs will also appear for the placement of the desired date.

In figure 5.2 is the login process if the user chooses automatic login.

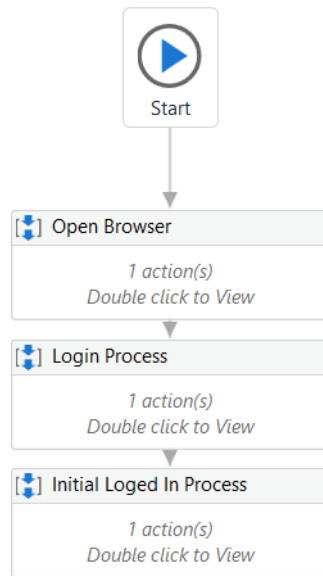


FIGURE 5.2: Automatic Login Process

This automatic login process is divided into three steps.

The first consists of opening the browser to be able to work with the banking platform, figure 5.3.

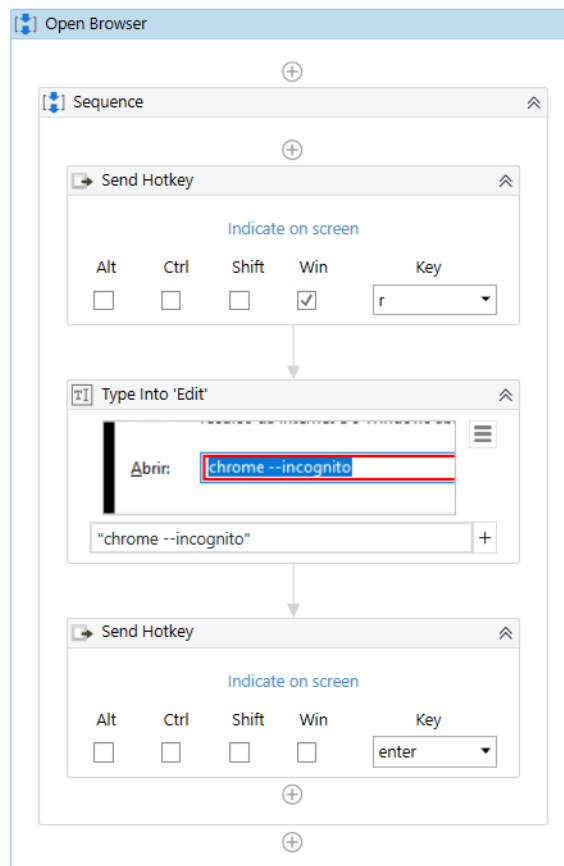


FIGURE 5.3: Open Browser Process

Analyzing in detail the process of opening the browser, consists of sending it to the system of the Win+r hotkey that allows the opening of "Execute" where it will be possible, in addition to opening the browser in incognito mode, it will also allow the analysis of the language used in the system so that the location of files changes depending on the language of the system, for example: in Portuguese, the location of a file in Google Drive would be "G:/Discos Partilhados" while in English it would be "G:/Shared Drives" so to avoid these cases the bot parses the language used in "Execute" to prevent future file location problems.

The browser is opened in incognito mode, in this case, the bot works in the best way since Google does not collect data from the process executed by the bot.

Finally, another hotkey is sent to the system that corresponds to Enter Key, this key will open chrome in incognito mode.

The second step consists of logging into the same banking platform, figure 5.4.

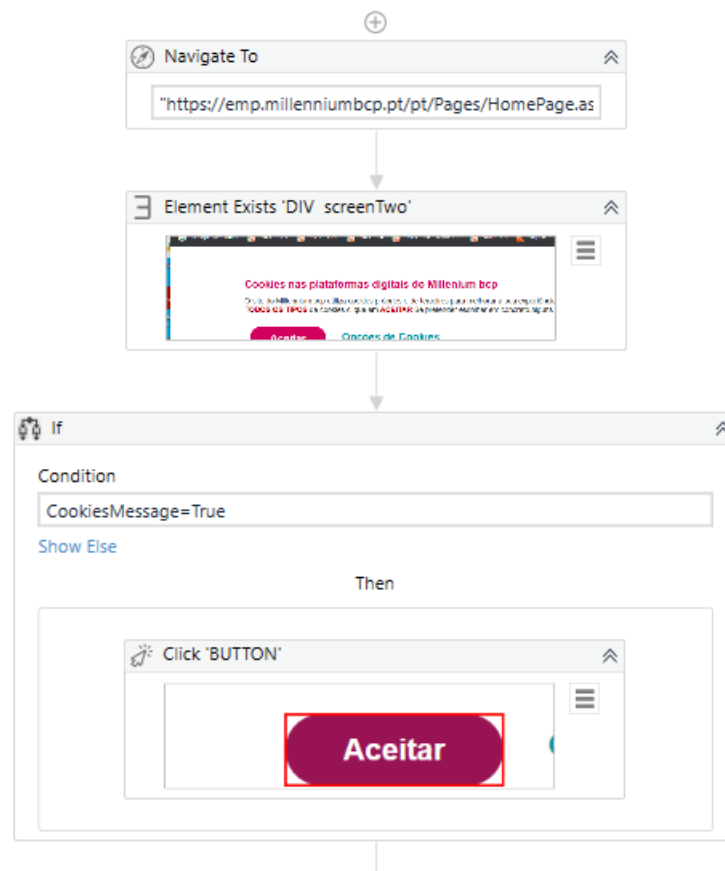


FIGURE 5.4: Navigate to page and accept cookies message.

In the figure, it is visible that, after the browser is open, it will navigate, in this particular case, to the Millennium BCP bank where the login process on the platform will begin.

Then the bot will wait for a Cookie message to appear, if it appears the bot will click on accept Cookies, if it does not appear the bot will proceed to the next task.

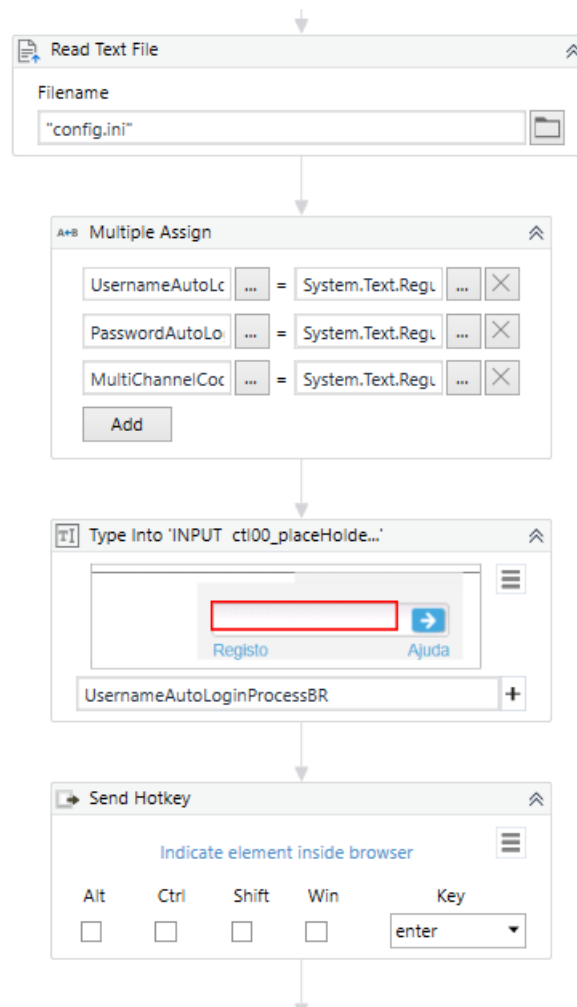


FIGURE 5.5: Read file and obtain login credentials

After accepting the Cookies in figure 5.5, the bot will read the "config.ini" file where it will find the necessary login data to log in to the banking platform. This data is obtained through regular expressions that will filter the intended data.

Then, the username obtained from the file is placed in the username input on the respective banking platform.

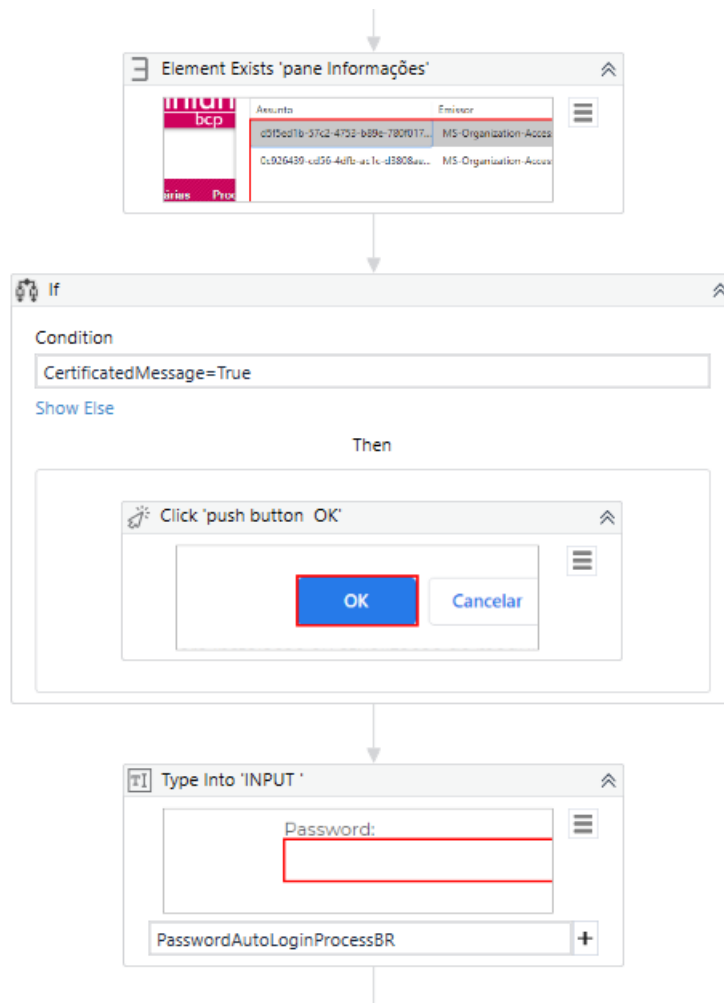


FIGURE 5.6: Certificates message and password insert.

Next in figure 5.6, the bot will analyze whether an element referring to the Millennium BCP certificates appears, if it appears it will press "OK". If it does not appear it will proceed to the next task.

Then the password value is entered in the password field of the banking platform.

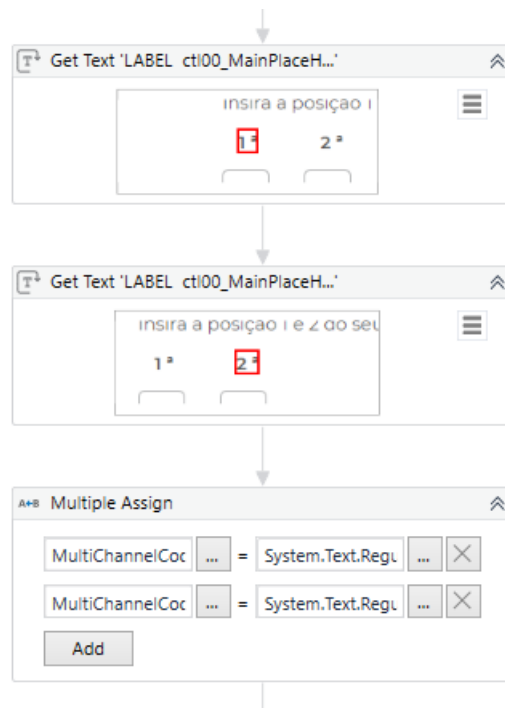


FIGURE 5.7: Insert multi-channel code.

Finally in figure 5.7, about the platform login process, the value of the position of the multi-channel code that needs to be inserted in the fields is read.

After reading the entire multi-channel code from the "config.ini" file, the bot will go through all the positions of that code and interconnect the position desired by the bank with the number present in the multi-channel code.

The before login steps are performed to download the desired documents.

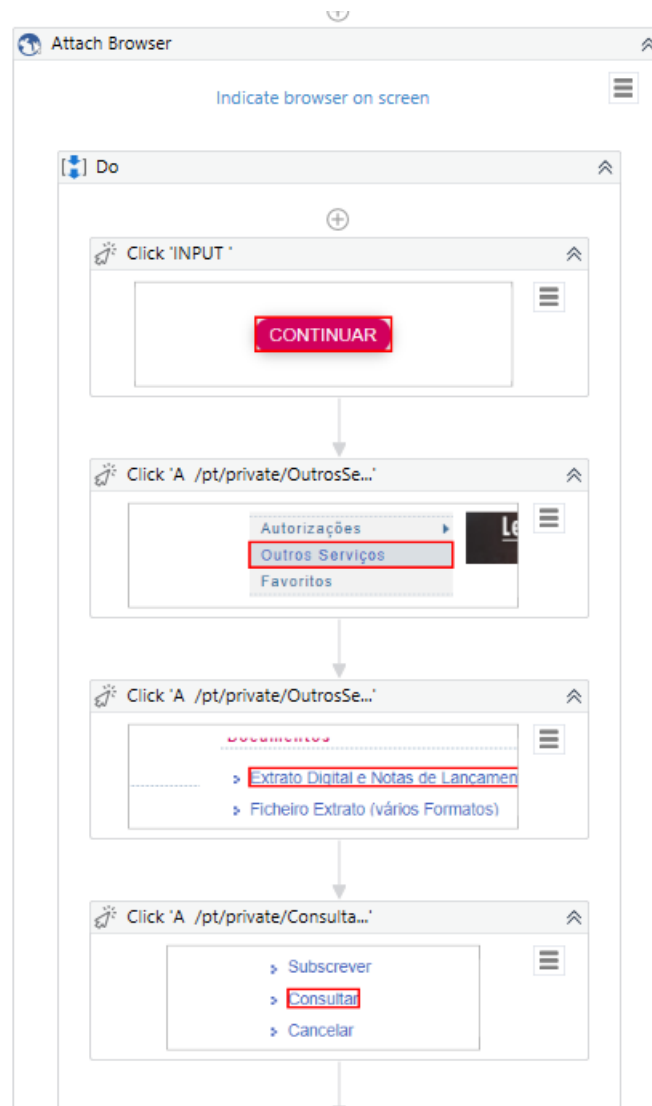


FIGURE 5.8: Click on the menus (first step).

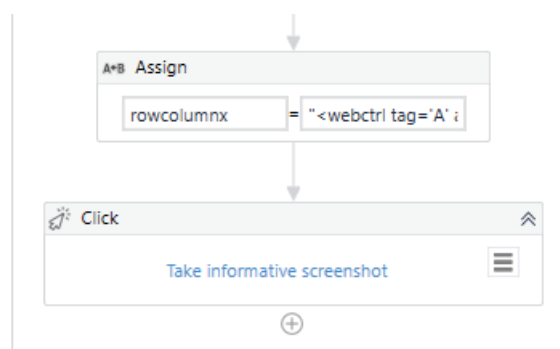


FIGURE 5.9: Click on the menus (second step).

In the last task related to the automatic login process, the bot will press the "continue" button to start the session and will go through the most diverse menus to download the invoice, figure 5.8 and 5.9.

The manual login process is the same as the automatic login process except that inputs appear to enter the user's login data on the banking platform.

Once the automatic and manual login process has been detailed, the manual date process is introduced in figure 5.10.

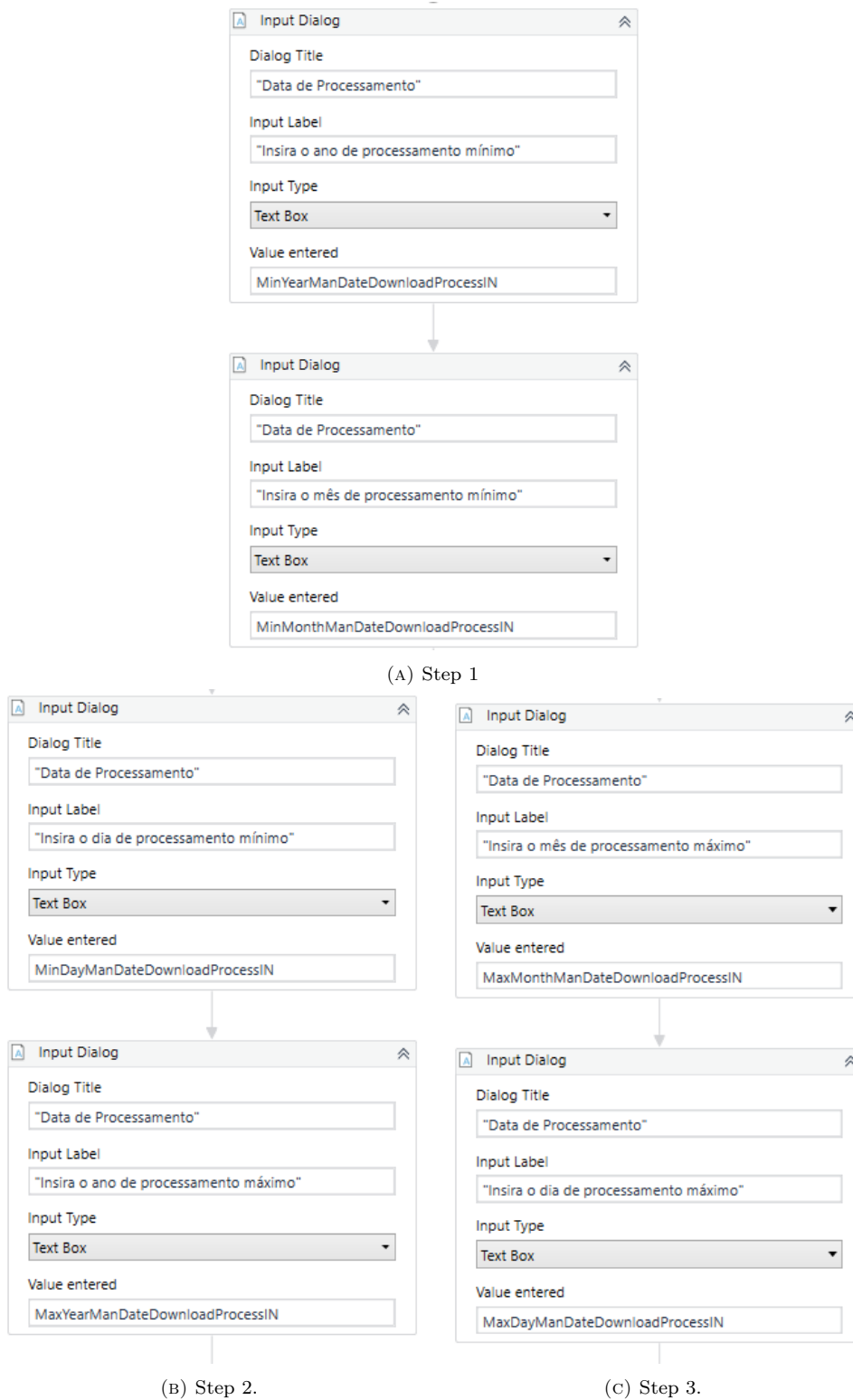


FIGURE 5.10: Manual Date Process

In the figure 5.10 referring to the manual date process, it is possible to observe that there is input data from the user, the values are saved in variables to be used later in the process.

The same happens in the automatic date process, the only difference is that there are no input fields of value by the user but a variable with yesterday's date is saved to be used in the process.

Once all the steps about the date process have been presented, the download process is now presented.

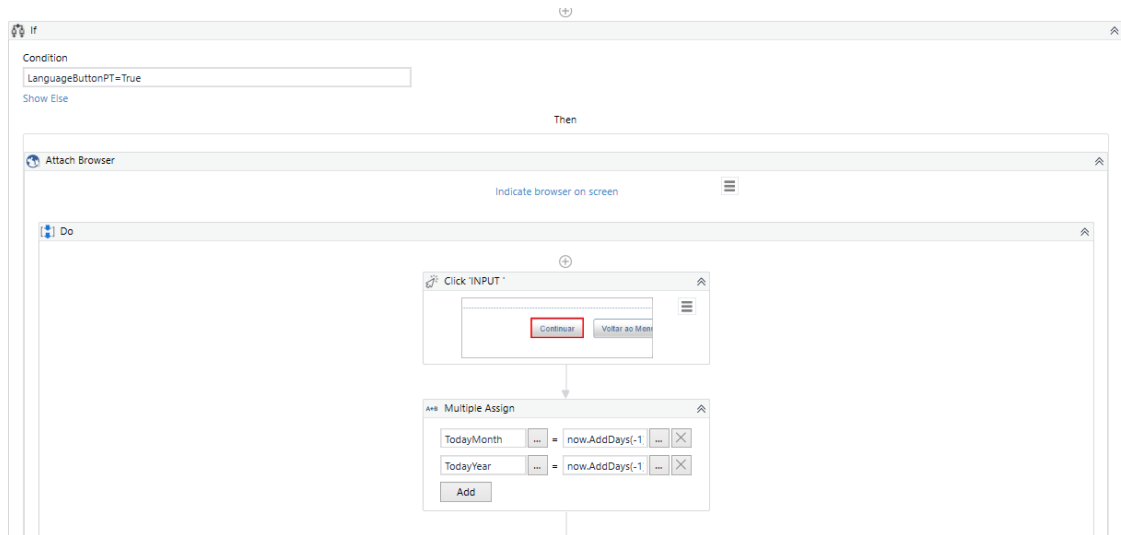


FIGURE 5.11: Create variables for yesterday's date.

In figure 5.11, an if is created where the language variable is used, which is the variable previously declared where a Boolean value is present and which defines the system language, that is, it identifies whether the language is Portuguese or English.

Afterward, the variables referring to the previous day are created since this download process in question uses the automatic date.

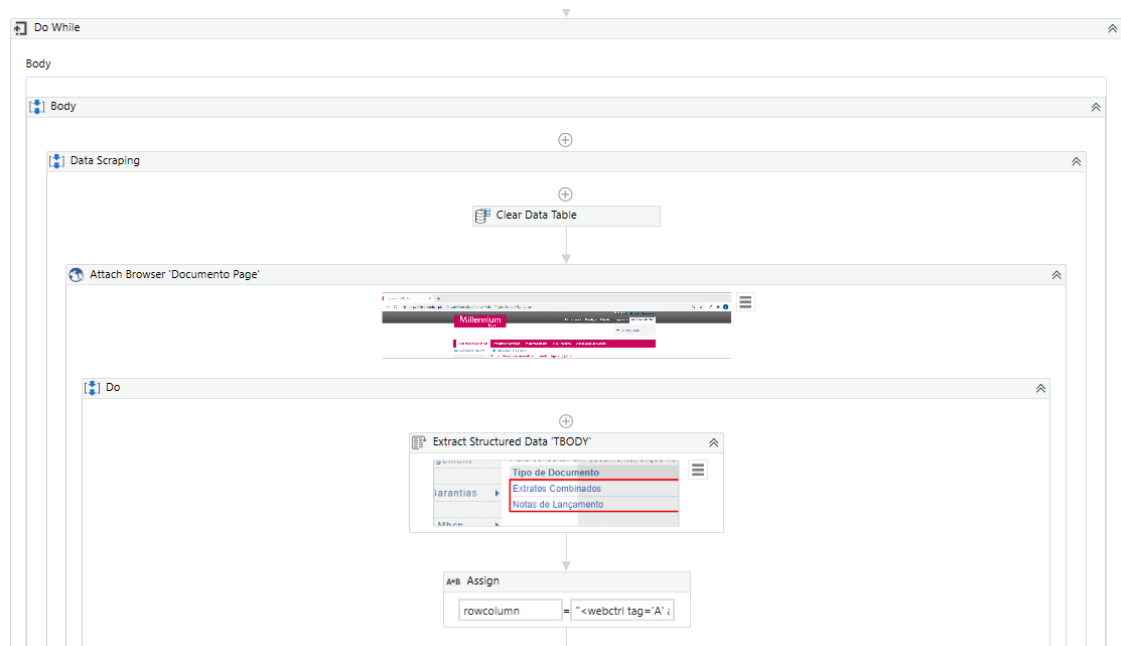


FIGURE 5.12: Click in the text release notes from the first table.

Within this table in figure 5.14, the objective is to click on all the links in the table, when clicking on it, the respective document is downloaded.

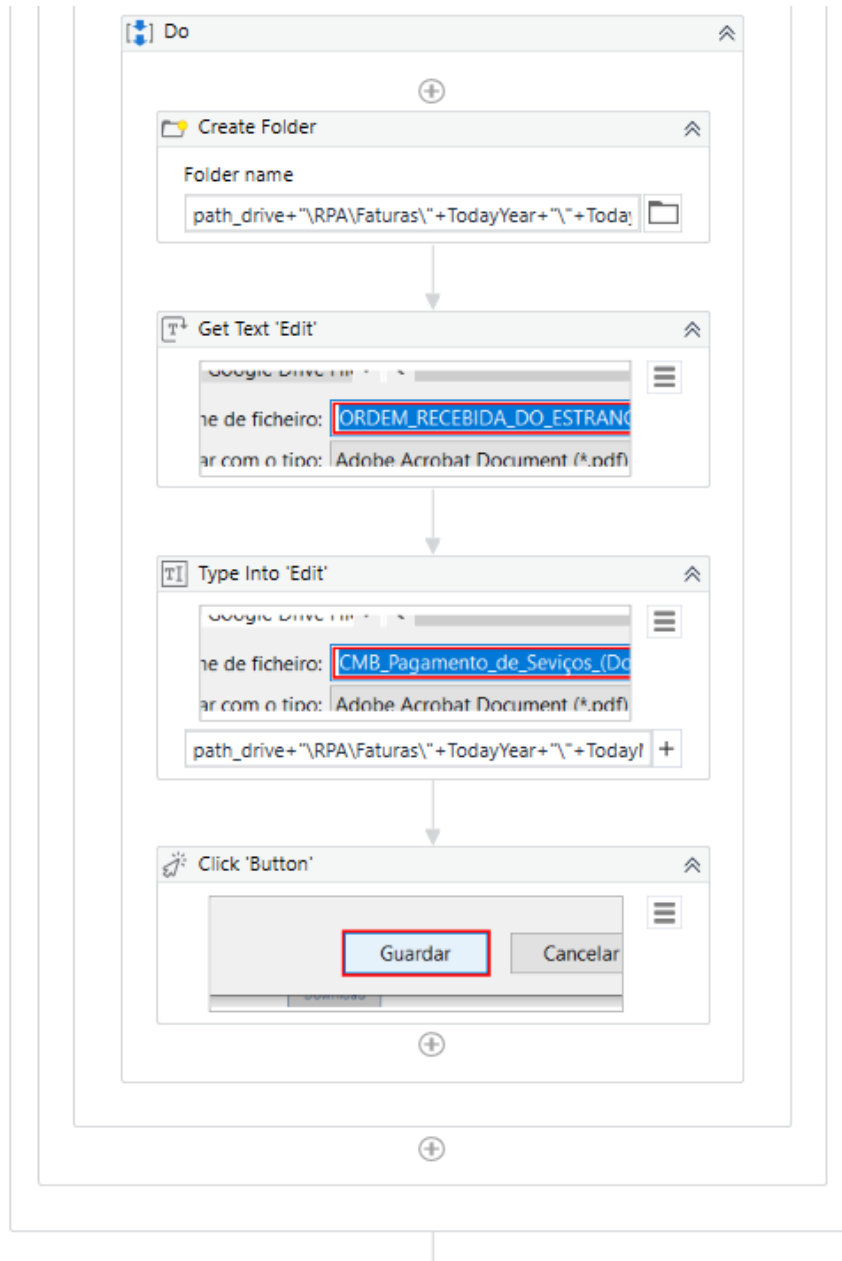


FIGURE 5.15: Save document.

After clicking on each of the links, a "save as..." box will appear in chrome where the path referring to that respective document will be inserted, figure 5.15.

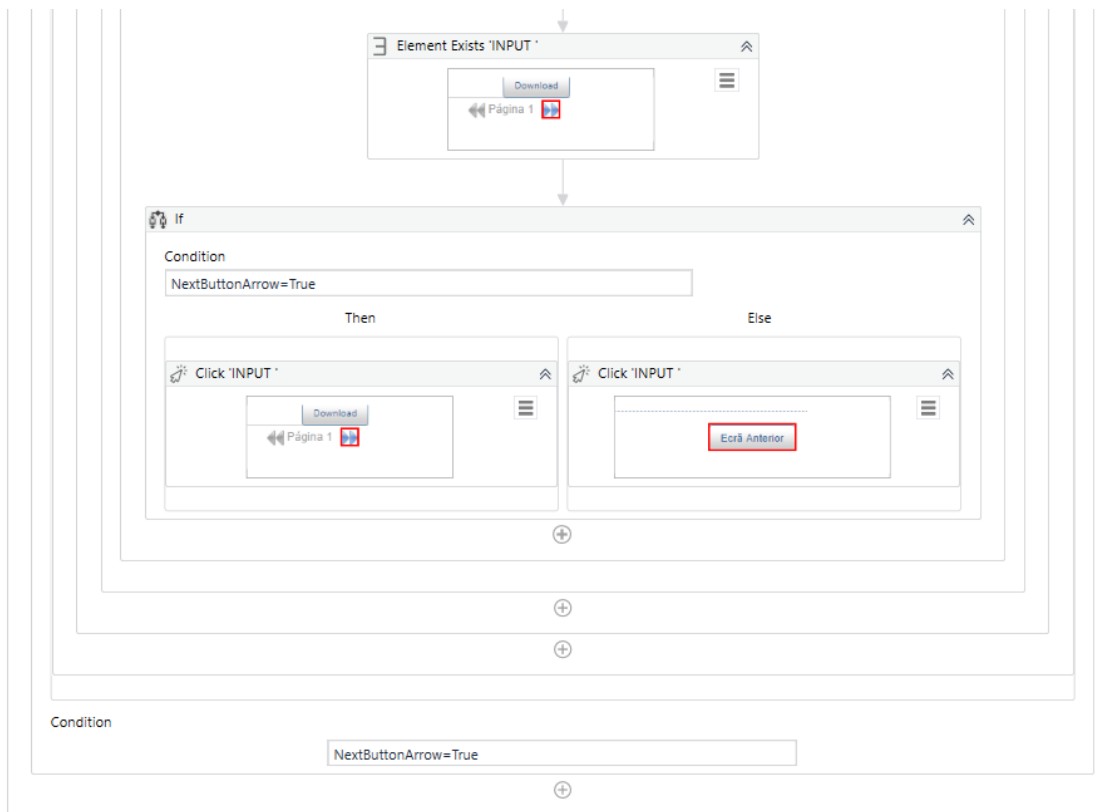


FIGURE 5.16: Check next button in data table.

After the first page in the table is finished, since the table is a data table and has several pages, the bot will see if there is some kind of button to go to the next page, if there is it advances if it does not exist, it returns since all documents have been downloaded, figure 5.16.

5.2 Extraction Process

The extraction process, like the download process, is a process common to both bank receipts and invoices.

This section depicts the extraction process for invoices. To perform this process, first, the invoices download for the month to be extracted must be completed.

As explained in subsection 4.2.6, this process regarding invoices aims to extract all the pages that contain the word 'Invoice-Receipts' from all Portable Document Format (PDF) previously downloaded so that they can be renumbered in the next process.

This process can be divided into two steps.

The first step concerns the initial inputs where the input data for the operation of the process are placed, such as the invoice extraction date.

In the second step, the documents are extracted for the dates introduced in the previous sub-process.

This extraction process aims to go through all previously downloaded documents and extract those that have the word 'Invoice-Receipt' to be renumbered in the renumbering process since only receipt-invoices can be renumbered.

The figure consists of two screenshots of a software dialog box titled "Input Dialog".

The top screenshot shows the following fields:

- Dialog Title: "Extraction"
- Input Label: "What is the intended extraction year?"
- Input Type: Text Box
- Value entered: ExtractionYearProcessIN

An arrow points from the top screenshot to the bottom screenshot.

The bottom screenshot shows the following fields:

- Dialog Title: "Extraction"
- Input Label: ""
- Input Type: Text Box
- Value entered: ExtractionMonthProcessIN

FIGURE 5.17: First step of the extraction process.

In figure 5.17 shows the inputs for both the month and the year in which the documents are to be downloaded.

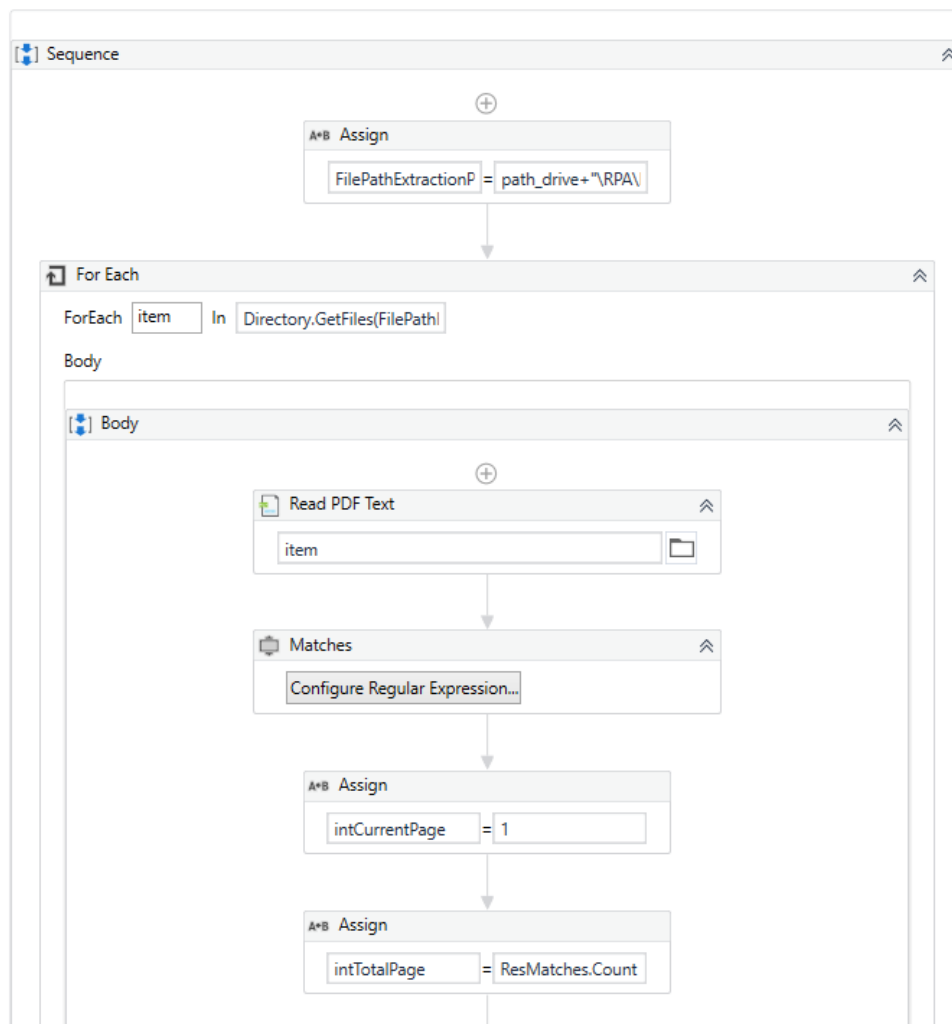


FIGURE 5.18: Start of the second step of the process.

In figure 5.18, it is possible to observe that a variable called 'FilePathExtractionProcess' is created, which aims to house the path to the PDF that will be used in this process.

Then, the cycle that will go through all the documents previously stored in the variable 'FilePathExtractionProcess' and apply the extraction process is started.

Afterward, the entire PDF document that is being used in the process is read and a matching case is made that aims to go through all the pages of the document to find the word 'Invoice-Receipt'.

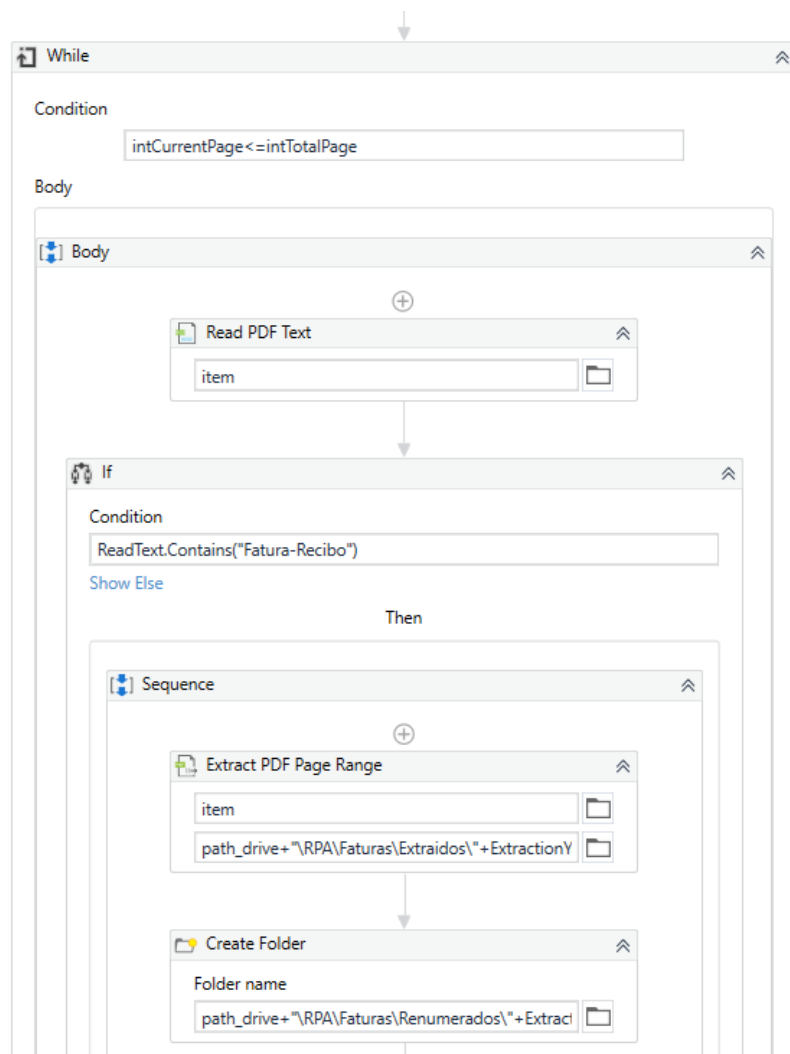


FIGURE 5.19: Middle of the second step of the extraction process

In figure 5.19, starts in a cycle that goes through all the pages of the PDF that is in the process. Then, the entire PDF is read and the pages that contain the word 'Invoice-Receipt' are searched for.

As soon as the pages containing the word 'Invoice-Receipt' are found, they are extracted into different files, which contain the page number and the name of the original file individually.

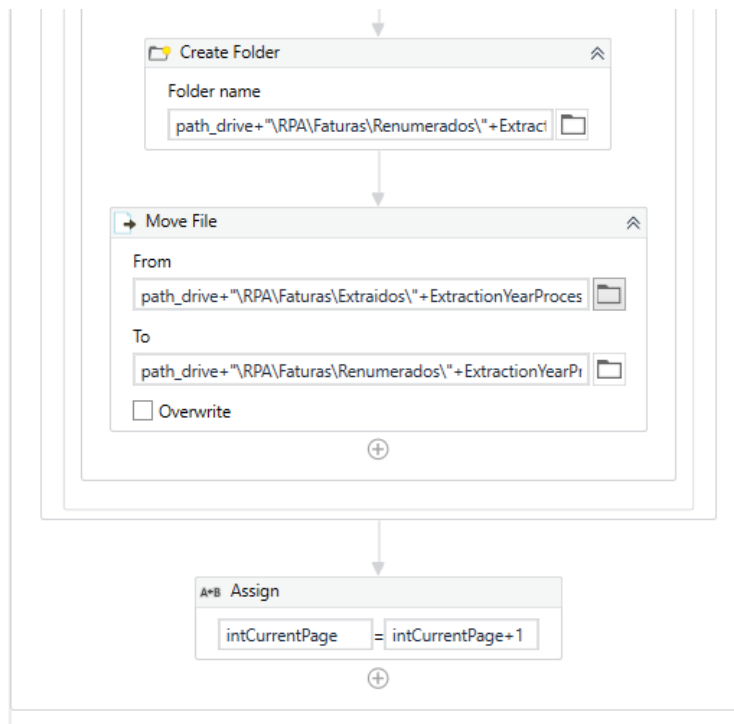


FIGURE 5.20: Final of the second step of the extraction process.

Finally in figure 5.20, the extracted pages are moved to a folder where they can be renumbered in the future.

5.3 Renumbering Process

The renumbering process aims to renumber invoices, more precisely the invoice receipt. For this process it was necessary to create a Python script to edit the PDFs.

This process should only be performed after downloads and extractions have been made for the months that are intended.

As an example, if it is necessary to renumber the files for May the download and extraction process for invoices related to May should have been carried out.

This process can be divided in two steps, one step consisting of the input fields for the date for which the bot should run and a second step where the bot will go through all the PDF and edit each one with the number received through the API.

As previously mentioned, the first step of the renumbering process consists of the date input, which refers to the date that the bot should execute, this must be given the year and the month of the date, for example for the bot execute one month the date should be 2021-05 for may 2021.

As for the second step of the process, it starts by reading all the PDF previously downloaded and extracted to be able to read the entry number that will return the internal number through the API.

For that, the bot needs to connect to the API, run the input query with the number that is read in each of the PDF and finally write that internal number in the PDF with the call of Python code.

This Python code will have to input two arguments which are in turn the location of the file in which the addition of the internal number is applied and the internal number itself.

This code is demonstrated on listing 5.1.

```

1 \\replace.py
2 from reportlab.pdfgen import canvas
3 from PyPDF2 import PdfFileWriter, PdfFileReader
4 from reportlab.lib.colors import magenta, red

```

```

5 import os, argparse, pikepdf
6 parser = argparse.ArgumentParser()
7 parser.add_argument('--inputfile')
8 parser.add_argument('--number')
9 args = parser.parse_args()
10 text = ('Number '+args.number)
11 file_path = (args.inputfile)
12 c = canvas.Canvas('watermark.pdf')
13 c.setFontSize(22)
14 c.setFont('Helvetica', 10)
15 c.setFillColor(red)
16 if ((len(text))<108):
17     xposition=433
18     if ((len(text))>26):
19         xposition=(433-(3.5*(len(text))))
20     c.drawString(xposition, 820,text)
21     c.save()
22     watermark_opener=(open("watermark.pdf", "rb"))
23     watermark = PdfFileReader(watermark_opener)
24     output_file = PdfFileWriter()
25     pdf = pikepdf.open(file_path, allow_overwriting_input=True)
26     file_unicrypted=(file_path.split('.pdf')[0]+'_unicrypted'+'.pdf')
27     pdf.save(file_unicrypted)
28     pdf.close()
29     input_file_opener=(open(file_unicrypted, "rb"))
30     input_file = PdfFileReader(input_file_opener)
31     page_count = input_file.getNumPages()
32     for page_number in range(page_count):
33         input_page = input_file.getPage(page_number)
34         input_page.mergePage(watermark.getPage(0))
35         output_file.addPage(input_page)
36     output_path = file_path.split('.pdf')[0] + '_' +args.number+ '.pdf'
37     with open(output_path, "wb") as outputStream:
38         output_file.write(outputStream)
39     input_file_opener.close()
40     watermark_opener.close()
41     os.remove(file_unicrypted)
42     os.remove(file_path)
43     os.remove(os.getcwd()+'/watermark.pdf')
44 else:
45     print("Error, the number is too large")

```

LISTING 5.1: Code for edit the PDF with internal number.

As it is visible in the code, libraries are used to renumber the PDF. Below, all the libraries used are detailed:

- **reportlab.pdfgen** - The pdfgen package is the lowest level interface for generating PDF documents.
A pdfgen program is essentially a sequence of instructions for "painting" a document onto a sequence of pages. The interface object which provides the painting operations is the pdfgen canvas.
- **PyPDF2** - The PyPDF2 package is a pure Python PDF library that you can use for splitting, merging, cropping, and transforming pages in your PDF.
- **os** - Python OS package allows us to use the operating system-dependent functionalities and to interact with the underlying operating system in several different ways.
- **argparse** - The argparse package makes it easy to write user-friendly command-line interfaces. The program defines what arguments it requires, and argparse will figure out how to parse those out of sys.argv.
- **pikepdf** -The pikepdf package is a Python library allowing creation, manipulation and repair of PDF.

After importing all libraries, the input arguments are created using the `argparse` library, which, as mentioned above, creates the two input arguments, the location and the number to be placed in the PDF.

After creating the Python code input arguments, a "text" variable is created with the text "Number" and the number received through the input argument.

Then, the variable "file_path" is created with the path of the PDF document to be edited and another variable is created too that will support the watermark that will be added to the PDF to be able to edit it.

In this watermark, the font size, color, and size of the text, as well as its position in the PDF, are defined.

Then a watermark will be placed on the PDF to be edited where the text to be added will be the number received by the argument of the code presented in 5.1.

At the end of this process, the newly edited document with the name of the previous PDF is saved.

To finish the whole process, the initial PDF is removed, as we create a new one that has already been edited and the file is encrypted to be able to read the entire document.

In the next figure presented are screenshots of this complete UiPath process.

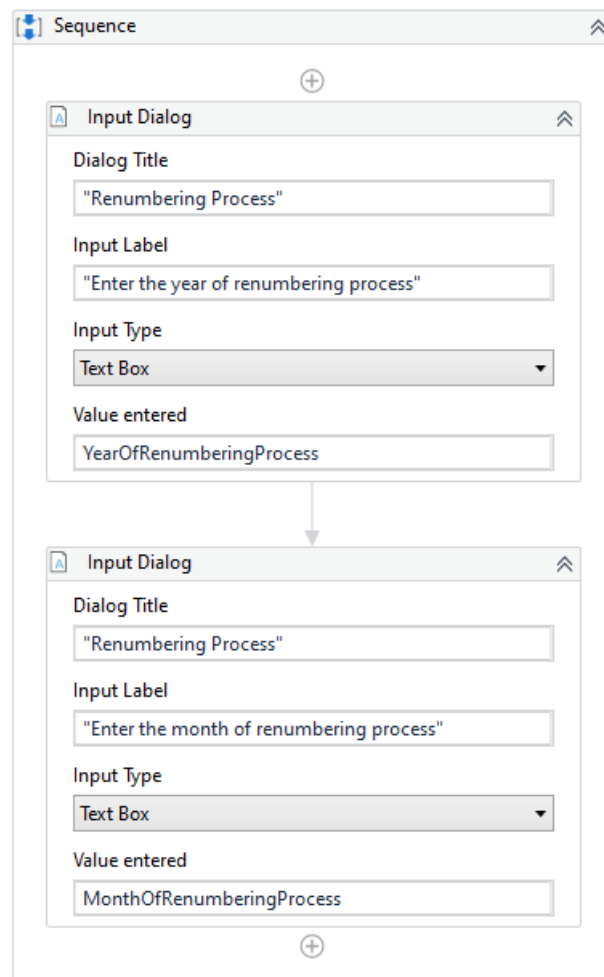


FIGURE 5.21: First step of the renumbering process.

In figure 5.21, it is possible to observe that two inputs are created, one for the year and the other for the month of the process.

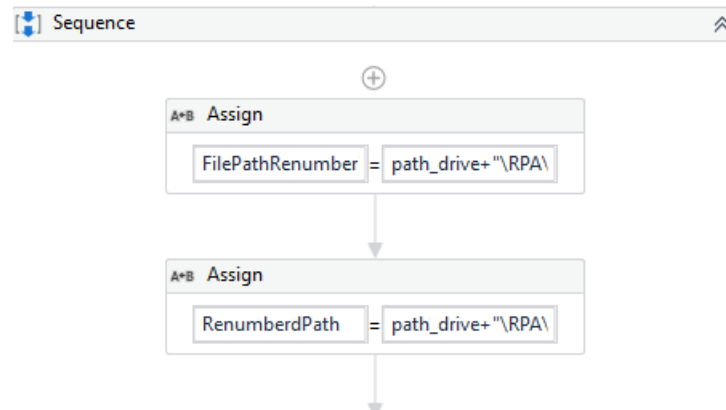


FIGURE 5.22: Creation of the file location variables.

In figure 5.22, it is possible to observe that two variables are created for the location of the file, one "FilePathRenumberingProcess" and the other "RenumberdPath".

The first variable indicates the location of the file to be renumbered and the second variable indicates where the document will be moved after being renumbered.

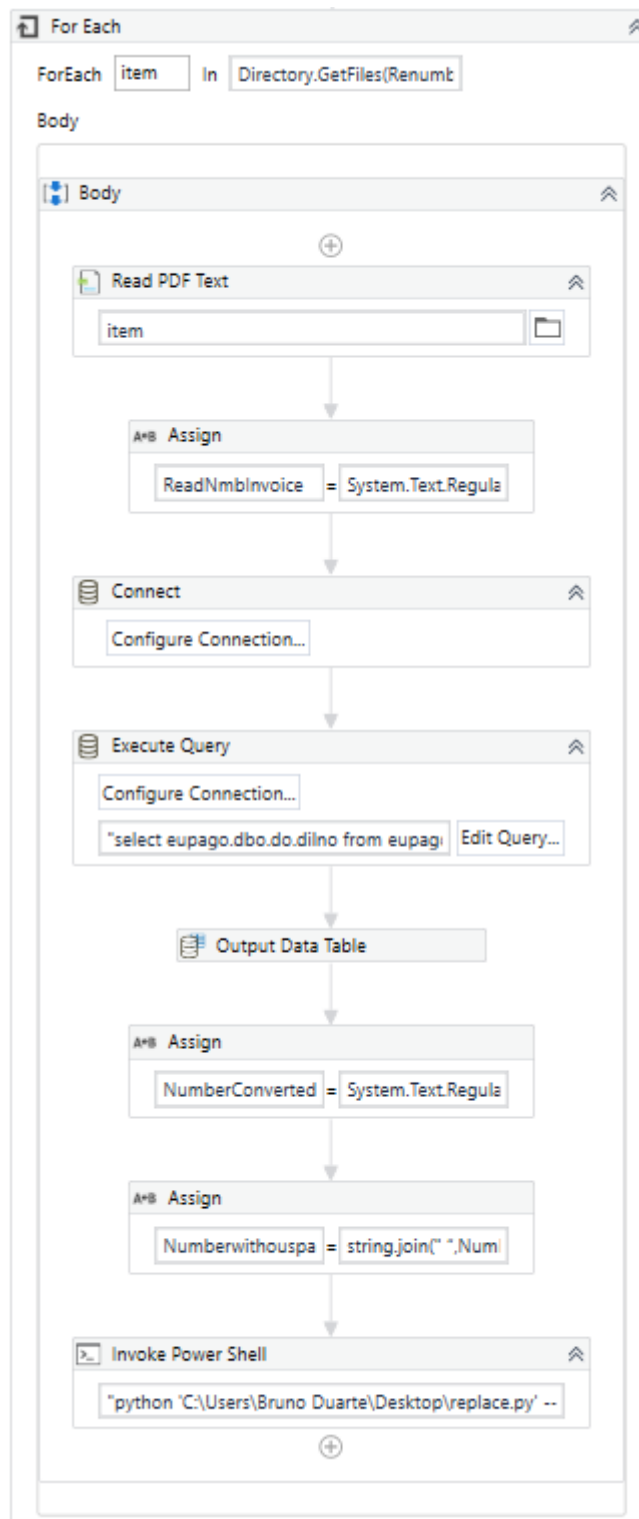


FIGURE 5.23: Second step of the renumbering process.

Finally in figure 5.23, the second step of the renumbering process is created.

In this last step, first a for loop is created that will execute this process for all PDF present in the variable "FilePathRenumberingProcess".

5.4 Print Process

In this section the print process is demonstrated. This demonstration follows the design drawn in the subsection 4.2.7.

The objective of this process is to join all documents from a certain month after being downloaded and renumbered in a single PDF and, finally, send them to the printer to be printed.

Below, the process in UiPath is presented.

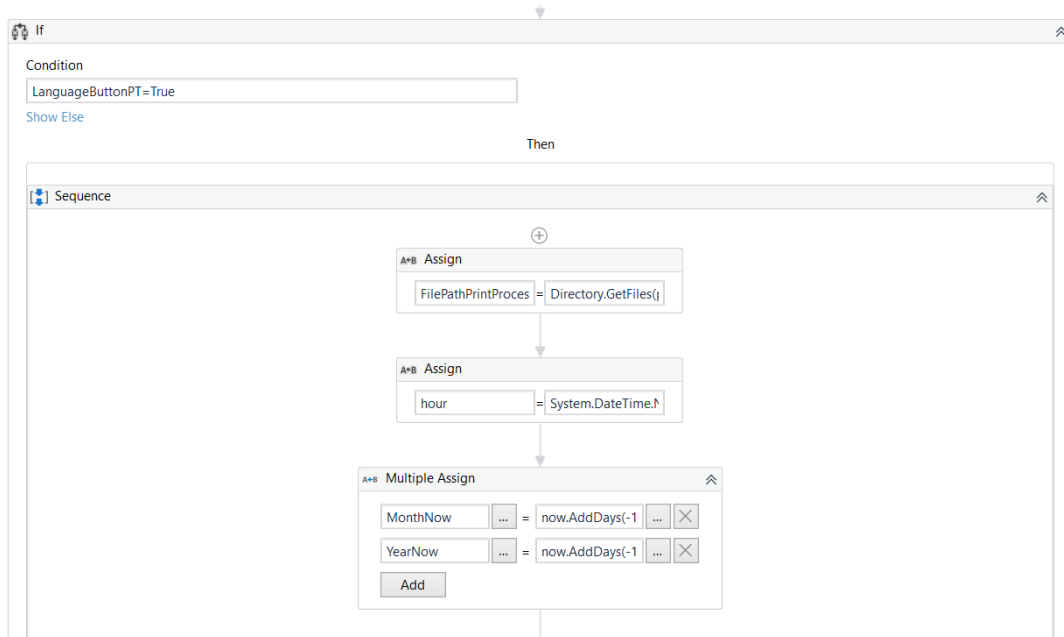


FIGURE 5.24: Create initial variables.

In the figure shown above, the creation of input variables referring to the month and year of printing the documents is shown.

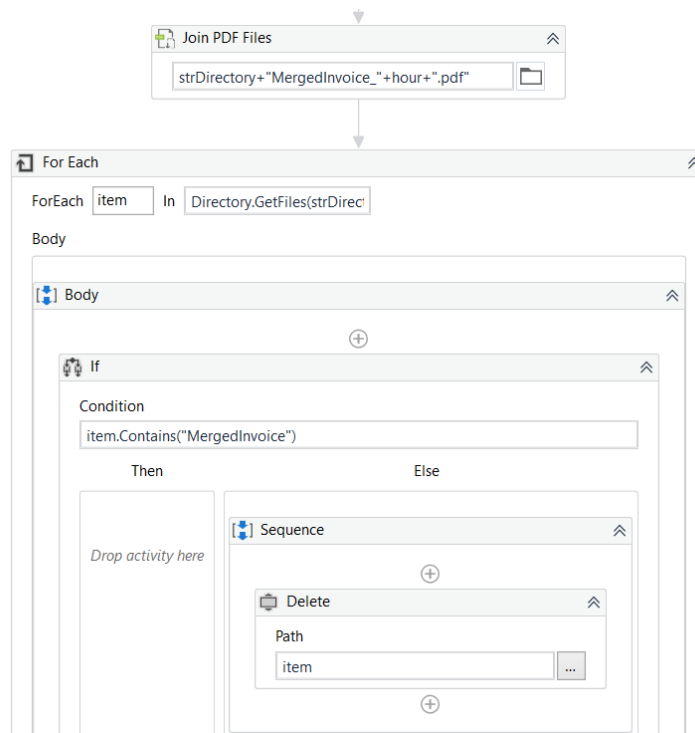


FIGURE 5.25: Join files.

After the creation of the initial variables, all the documents presented under the folder referring to the entered month and year are joined.

Once the document with all the renumbered files is finished, the bot will analyze that folder, if it contains any files besides the joined file it will delete it so as not to repeat files in the future.

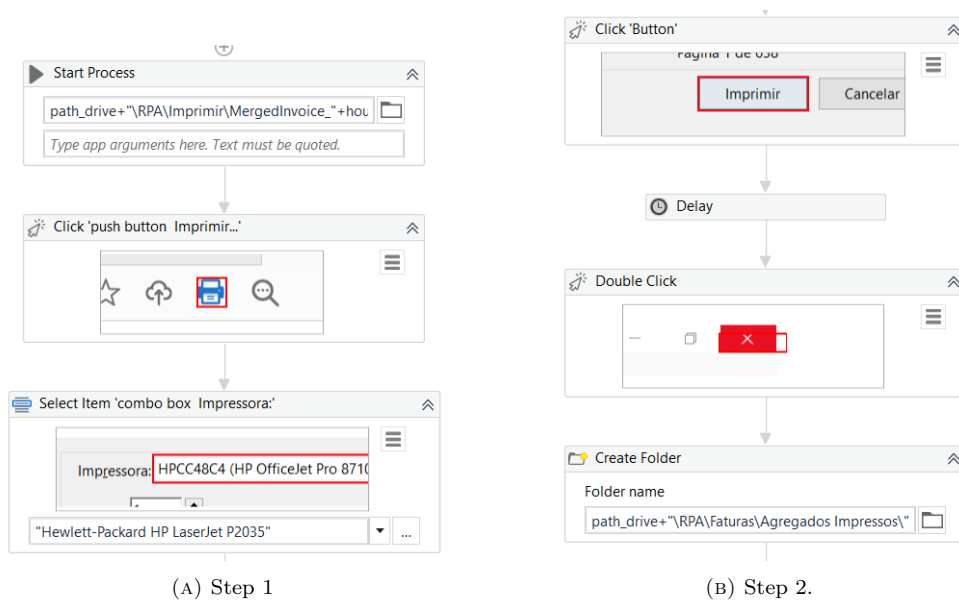


FIGURE 5.26: Send file to printer.

After joining all documents into one, this file is opened and sent to the respective printer.

All these processes presented in this section work both in isolation and together, that is, the user can start only one of these processes or choose several and start them.

The bot will sequentially follow the processes, for the, for example, renumbering process to work, both download and extraction processes must be performed.

Chapter 6

Evaluation

This chapter aims to define the evaluation procedure that is implemented as a way to evaluate the solution that is built for the existing problem. For that, it will be necessary to use a set of parameters so that it is possible to correctly evaluate the solution.

6.1 Metrics

To be able to evaluate this solution, the following metrics are used:

- Performance (percentage);
- Task execution time (minutes);
- Error rate (numeric value);
- Number of files affected (numeric value);
- Bounce rate (percentage);
- Employee satisfaction (numeric value).

In the following sections, each of these metrics that are used to evaluate the solution is explained in detail.

6.1.1 Performance

This metric aims to analyze the performance of both the bot in each of the tasks it performs as well as of the employee so that it is possible to evaluate which one obtains the best performance rates for each of the tasks.

With this metric, it will be possible to understand the percentage of gain in the execution of the largest number of tasks in the minimum possible time.

6.1.2 Task execution time

The task execution time metric is very similar to performance, but in this metric, the objective is something more specific like the execution time that both the bot and the employees spend for each of the tasks performed.

6.1.3 Error rate

This metric, as the name implies, aims to analyze the rate of errors that are made by the bot and the employees.

6.1.4 Number of files affected

This metric aims to analyze the number of files affected by both the bot and the employees.

Each of these files is made up of several documents, which may cause their size to change.

6.1.5 Bounce rate

The bounce rate metric aims to analyze the rejection rate, by employees, over the time of using this type of solution so that in this way it was possible to analyze the next metric.

6.1.6 Employee satisfaction

This metric aims to analyze the employee satisfaction rate in using this type of solution.

In conclusion, with these metrics, it is intended to obtain the necessary data to evaluate the solution to understand the impact it has on a company.

6.2 Hypotheses

In this section, all the hypotheses that are evaluated in this solution will be presented.

- H1. The bot performs better than a human user;
- H2. The bot has a shorter task execution time than a human user;
- H3. The bot has the same error rate as a human user;
- H4. A human user has fewer files affected per minute than a bot;
- H5. A human user has a higher bounce rate when using a bot;
- H6. A human user has a higher rate of satisfaction with the use of a bot.

Thus defined the hypotheses to be evaluated, it is possible to observe that there are some conservative hypotheses as is the fact of H4 since there will be no difference with the implementation of this solution.

About the other hypotheses, they are all presented as alternative hypotheses since they change with the implementation of the solution

6.3 Evaluation methodology

To evaluate the solution based on the previously mentioned metrics, it is necessary to implement a methodology.

This methodology involves the use of survey results, analysis of data previously obtained about the time spent and number of files affected for each of the tasks performed by employees.

Finally, a control/test group is also used to be able to evaluate this solution.

6.3.1 Survey

A 5 Point Likert Scale will be used in order to be able to perceive and answer certain questions that will appear in the same.

Below are examples of questions that will appear in this survey.

- How easy is the solution to run?
- How easy is it to understand at what point in the process is the solution?
- What is the level of error rate?
- What is the total time level of the processes?

6.3.2 Control/Test Group

This control group will consist of three people who perform financial functions in the company.

These people are the ones who carry out these types of processes manually, so now it is intended to put them to use this solution to be able to do tests.

6.4 Data analysis

The data obtained through samples taken from the processes executed manually and through the implemented solution are compared.

The data presented in the next tables refers to processes executed manually, after these data are presented, they are compared with data referring to the same processes for the same period and with the same number of documents through the UiPath solution.

TABLE 6.1: Number of invoices download in last 3 months.

Month	Number of files
March	2320
April	2167
May	2122

In table 6.1 it is possible to view all documents downloaded for the corresponding months.

TABLE 6.2: Time for download fifty invoices.

Month	Time Manually	Time Automatically
March	21 min	8 min
April	19 min	5 min
May	17 min	4 min
Average	18,6 min	5,6 min

In table 6.2, it is possible to view the download times for fifty files either through the RPA process or through the manual process.

By observing these values it is visible that there is a large discrepancy between the download times referring to the RPA process and the manual process.

The average time taken by the manual process is about three times longer than the process through RPA.

Are shown the values for the two different types of processes but in this case for bank receipts.

TABLE 6.3: Number of bank receipt download in last 3 months.

Month	Number of files
March	769
April	663
May	640

The table above shows the total number of files referring to bank receipts.

TABLE 6.4: Time for download fifty bank receipts.

Month	Time Manually	Time Automatically
March	1h03 min	28 min
April	1h21 min	34 min
May	1h07 min	28 min
Average	1h10 min	30 min

In the table it can be seen that in the case of bank statements, the process with the RPA is about two times faster than the process without the RPA.

In the case of the other three processes, they were never carried out manually as they are very time-consuming and specific processes, which could not be done without the help of the RPA.

With no basis for comparing manual processes and processes with the RPA as was done for the download processes, the numbers of these processes will only be shown with the help of the RPA.

Through the analysis of tables 6.2 and 6.4, it is possible to prove the hypotheses H1, H2 and the hypothesis H4 is contradicted.

TABLE 6.5: Extraction process with RPA solution.

Month	Time
March	4 min
April	4 min
May	3 min
Average	3,66 min

TABLE 6.6: Renumbering process with RPA solution.

Month	Time
March	6 min
April	4 min
May	4 min
Average	4,66 min

In the two tables presented, it is visible to see that these processes carried out by the RPA are quite fast.

6.5 Survey Results

In this section, the results obtained in the survey carried out are shown. This survey was answered by the people involved in the test group.

The scale of each of these questions is from 0-5 where 0 is easy/low and 5 is difficult/high.

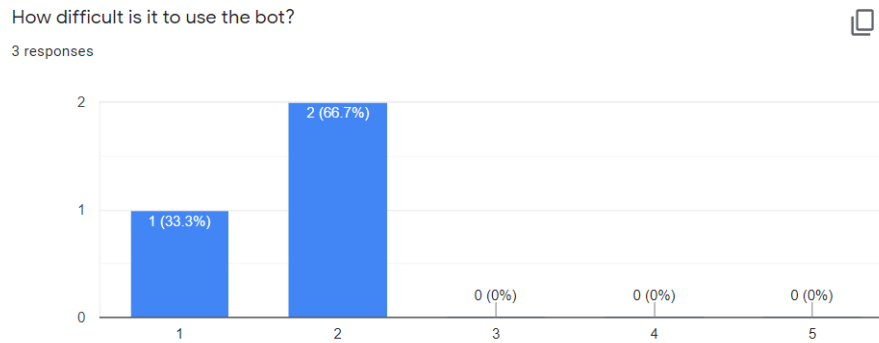


FIGURE 6.1: Answer to first question.

How easy is it to understand at what point in the process is the solution?

3 responses

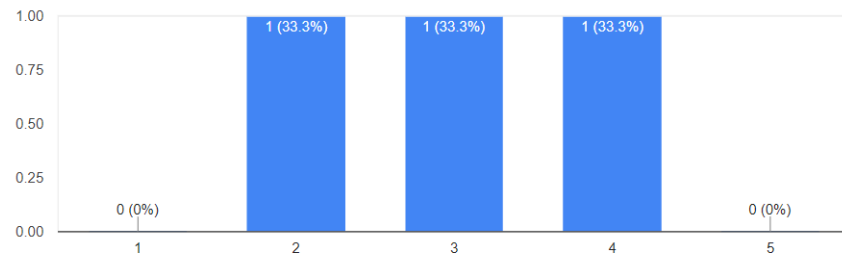


FIGURE 6.2: Answer to second question.

What is the level of error rate?

3 responses

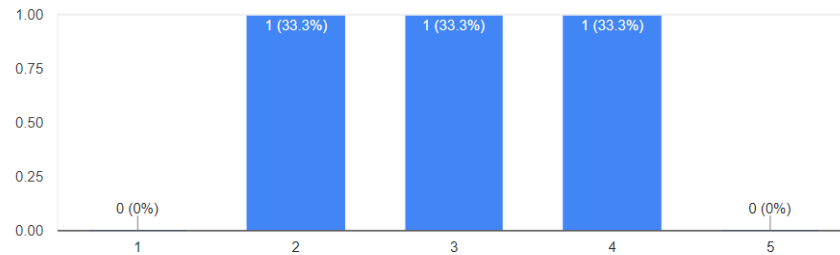


FIGURE 6.3: Answer to third question.

What is the total time level of the processes?

3 responses

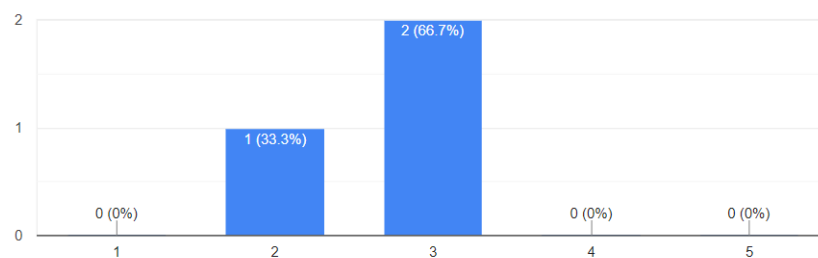


FIGURE 6.4: Answer to fourth question.

The survey data vary a lot, this situation can be explained with the perception and the greater use of technologies from one another about the other, as they are from different age groups.

In this way, it is possible to obtain more realistic data given the difference that exists between the answers and thus obtain more concrete data.

Analyzing the figures 6.1 and 6.2, it is possible to observe that hypotheses H5 and H6 are proven.

In summary, when analyzing this data, it is possible to see that the interviewees think that this solution is very easy to use and the time for each process is relatively low.

Chapter 7

Conclusion

In conclusion, the results obtained using a manual procedures vs. an automatic procedure show a significant difference.

This same difference can be explained both by the fact that these processes are too boring and tiring for the employees, since they are repetitive and sequential processes, as well as by the fact that a process done manually is more prone to possible errors.

Therefore, with the proposed solution, these same problems were reduced.

These data also influence the bank's server as the times vary depending on the response from the bank, the idea is obtained since tests were performed both during the day, where there is a lot of influx of bank customers, as at night, where the influx is very low, and the values were quite different since the server had a much faster response and therefore the downloads turned out to be fast, this difference in the obtained values is not significant since the tests and values shown in data analysis section were obtained under the same circumstances.

In short, the solution greatly improved the processes that already existed in terms of time and the level of errors that existed in manual processes, and processes that were difficult to be done by a human user were added due to their difficulty or delay as it is the case of extraction and renumbering processes.

Also, when analyzing the data from the survey responses, it was noticeable that the people surveyed and involved in this process find this solution quite easy to use and that it takes relatively little time to implement.

The objectives defined at the beginning of this project were all successfully achieved this work can still be improved in the future as this development process requires a constant update of the interfaces it interacts with, for example, on bank websites, as change the interface it will be necessary to update the RPA process.

Therefore, this tool also becomes somewhat limiting as it requires interaction with the interface itself, which means that if there is any change in it, the process has to be updated.

Also, in the future, other processes at a financial level may be built using this tool to reduce the number of sequential and repeated processes in this area.

Finally, this solution fulfilled the initial objective of improving existing processes, decreasing the error rate, and freeing employees for more complicated tasks to be performed.

Bibliography

- [1] Nick Rich and Matthias Holweg. “Rich, BSc MBA Matthias Holweg, Dipl.- Wirtschafts-ing.(FH) MSc”. In: (2000), p. 32. URL: https://www.urenio.org/tools/en/value%7B%5C_%7Danalysis.pdf.
- [2] Alexander OSTERWALDER. “Designing a decision support system for handling heterogeneous spatial big data in resilient cities: Experiences from disaster risk management -”. In: *CWL Publishing Enterprises, Inc., Madison* 2004 (2004), p. 352. URL: http://biblioteca.usac.edu.gt/tesis/08/08%7B%5C_%7D2469%7B%5C_%7DC.pdf.
- [3] Peter A. Koen, Heidi M.J. Bertels, and Elko J. Kleinschmidt. “Managing the front end of innovation-part II: Results from a three-year study”. In: *Research Technology Management* 57.3 (2014), pp. 25–35. ISSN: 19300166. DOI: 10.5437/08956308X5703199.
- [4] Blue Prism. “Blue Prism Software Robots . The Art of Possibility .” 2016. URL: <https://www.blueprism.com/uploads/resources/white-papers/Blue-Prism-Product-Overview-Enterprise-Edition.pdf>.
- [5] Xavier Lhuer. “The next acronym you need to know about: RPA (robotic process automation)”. In: *Digital McKinsey* 17.December (2016), pp. 1–5. URL: https://www.mckinsey.com/business-functions/digital-mckinsey/our-insights/the-next-acronym-you-need-to-know-about-rpa%7B%5C_%7DOAhttps://usblearn.belpark.sun.ac.za/pluginfile.php/40119/mod%7B%5C_%7Dresource/content/0/2018MPHilFS%7B%5C_%7DRobotic%7B%5C_%7Dprocess%7B%5C_%7Dautomation%7B%5C_%7DMcKinsey%7B%5C_%7DCompany..
- [6] Wil M.P. van der Aalst, Martin Bichler, and Armin Heinzl. “Robotic Process Automation”. In: *Business and Information Systems Engineering* 60.4 (2018), pp. 269–272. ISSN: 18670202. DOI: 10.1007/s12599-018-0542-4. URL: <https://doi.org/10.1007/s12599-018-0542-4>.
- [7] Alok Mani Tripathi. *Learning Robotic Process Automation*. 2018. ISBN: 978-1-78847-094-0. URL: https://books.google.com.br/books?id=SLZTDwAAQBAJ%7B%5C_%7Dprintsec=copyright%7B%5C_%7Dhl=pt-PT%7B%5C_%7Dsource=gbs%7B%5C_%7Dpub%7B%5C_%7Dinfo%7B%5C_%7Dr%7B%5C_%7Dv=onepage%7B%5C_%7Dq%7B%5C_%7Df=false.
- [8] Lim Mei Ying. *Robotic process automation with Blue Prism quick start guide: create software robots and automate business processes*. 2018, p. 232. ISBN: 978-1-78961-446-6. URL: https://books.google.com.br/books?id=kC59DwAAQBAJ%7B%5C_%7Dprintsec=copyright%7B%5C_%7Dhl=pt-PT%7B%5C_%7Dsource=gbs%7B%5C_%7Dpub%7B%5C_%7Dinfo%7B%5C_%7Dr%7B%5C_%7Dv=onepage%7B%5C_%7Dq%7B%5C_%7Df=true.
- [9] UJJWAL GUPTA. “RPA AUTOMATION ANYWHERE”. Himachal Pradesh, 2019. URL: <http://www.ir.juit.ac.in:8080/jspui/bitstream/123456789/22756/1/RPA%20AUTOMATION%20ANYWHERE.pdf>.
- [10] Dylan Bowman. *RPA vs API Integration: How to Choose Your Automation Technologies*. 2020. URL: <https://www.uipath.com/blog/rpa-vs-api-integration>.
- [11] Aaron Guilmette. *Workflow Automation with Microsoft Power Automate*. Birmingham: Packt Publishing Ltd., 2020, p. 276. ISBN: 978-1-83921-379-3.

Appendix A

Value Proposition Canvas

In the next page, have a full page view of the value proposition canvas.

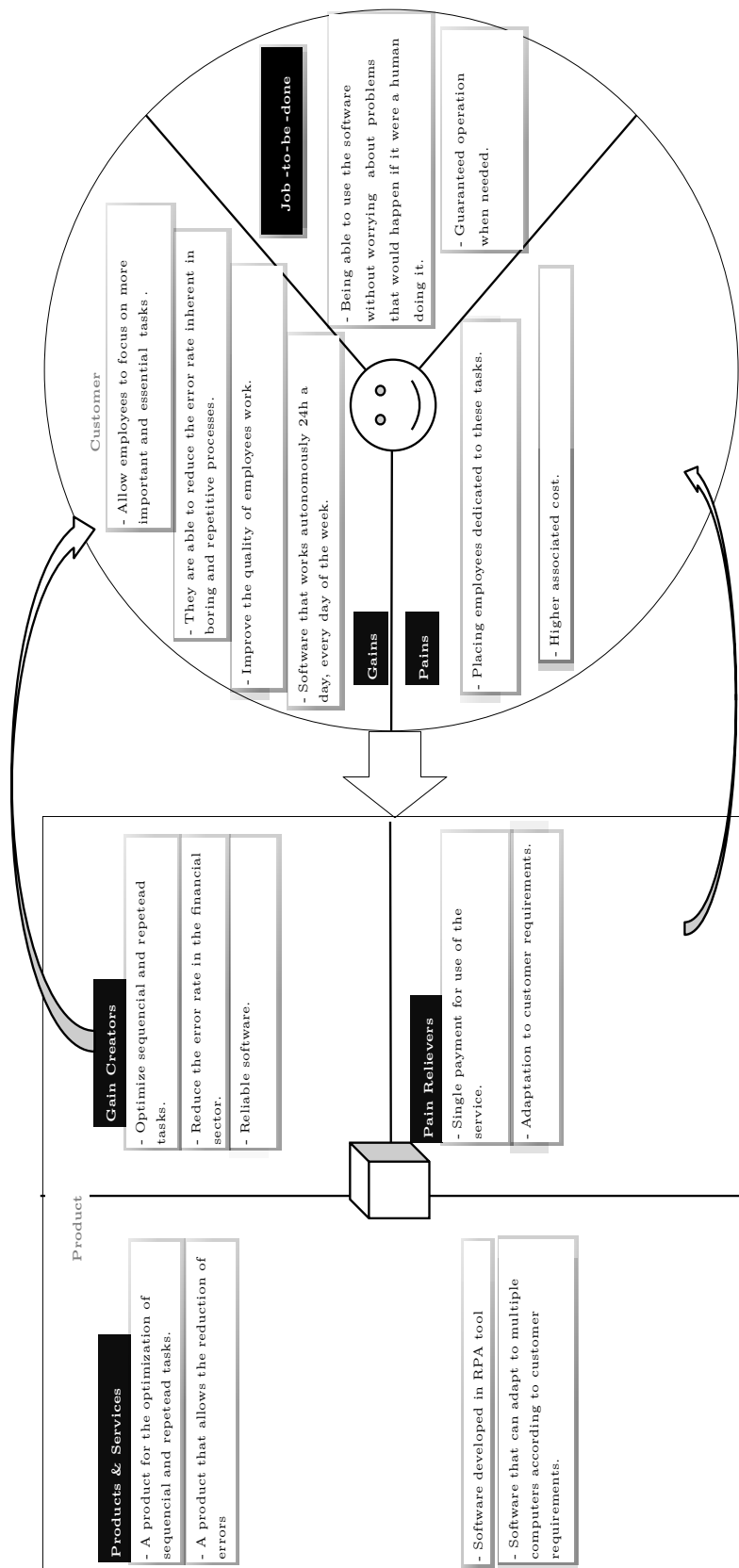


FIGURE A.1: Value Proposition - Canvas (Maximized)