

Impact of Robotic Process Automation in Accounting Industry

MSc Research Project

MSc in Fintech

Veena Hanumanthu Chenneti

x21170487@student.ncirl.ie

School of Computing

National College of Ireland

Supervisor: Brian Bryne

MSc Project Submission Sheet

School of Computing

Student Name: ...Veena Hanumanthu
Chenneti.....

Student ID: ...x21170487.....

Year: ...2022-23..

Programme: ...MSc in Fintech.....

Module: ... **Research Project (MSCFTD1)**.....

Supervisor: ...Brian Byrne.....

Submission Due Date: ...14 Aug 2023.....

Project Title: ... Impact of Robotic Process Automation in Accounting

Word Count:8092..... **Page Count**.....26.....

I hereby certify that the information contained in this (my submission) is information pertaining to research I conducted for this project. All information other than my own contribution will be fully referenced and listed in the relevant bibliography section at the rear of the project.

ALL internet material must be referenced in the bibliography section. Students are required to use the Referencing Standard specified in the report template. To use other author's written or electronic work is illegal (plagiarism) and may result in disciplinary action.

SignatureVeena Chenneti.....
:

Date:

PLEASE READ THE FOLLOWING INSTRUCTIONS AND CHECKLIST

Attach a completed copy of this sheet to each project (including multiple copies)	<input type="checkbox"/>
Attach a Moodle submission receipt of the online project submission, to each project (including multiple copies).	<input type="checkbox"/>
You must ensure that you retain a HARD COPY of the project, both for your own reference and in case a project is lost or mislaid. It is not sufficient to keep a copy on computer.	<input type="checkbox"/>

Assignments that are submitted to the Programme Coordinator Office must be placed into the assignment box located outside the office.

Office Use Only	
Signature:	
Date:	
Penalty Applied (if applicable):	

Impact of Robotic Process Automation in Accounting Industry

Veena H. Chenneti
Student ID - x21170487

Abstract

Robotic Process Automation (RPA) usage is rising along with technological advancements and causing a swift growth in its adaptation. It is a rapidly growing field which involves setting up of bots or software processes which can be used in accounting to save time for important tasks, focusing on its potential and challenges, especially in typical accounting scenarios involving repetitive tasks. In this research, the researcher focuses on the PoC (proof-of-concept) of RPA approach, which involves testing automation using sample invoices to compare human efforts with automation, processed through three different methods. Invoice booking manually involves entering details like item names, prices, and quantities one by one into a ledger or software. Microsoft power automate method involves creating a flow and Python method has use of libraries like pandas, openpyxl, pdfplumber. The researcher examines how it requires careful attention to avoid errors and ensure accurate record-keeping whereas automation is a one-time setup of processes. The comparison and evaluation is recorded with different manual efforts, using Microsoft power automate, and python libraries.

Keywords – Robotic process automation, pdfplumber, python, power automate, openpyxl,

Introduction

The term Robotic Process Automation (RPA) is a combination of robotics which is not an actual physical robot but a software which acts like an agent instead of human beings and processes the automation workflow into the systems (Syed et al., 2020). Despite RPA coming into prominence in the early 2000s, its initial development had already begun in the early 1990s. 'Machine Learning (ML)' stands among the technologies that contributed significantly to innovation, ultimately giving rise to the creation of RPA. As early as 1959, 'Arthur Samuel' played a pivotal role in developing Machine Learning. Machine Learning revolutionized computers' capabilities, enabling them to undertake essential tasks like translation and text summarization. However, there were inherent limitations in processing language. This led to the emergence of 'Natural Language Processing (NLP),' a technology that enhanced computers' ability to comprehend and process human language more accurately. In 1960, NLP amalgamated with 'AI (Artificial Intelligence)' to facilitate interactions between computers and human languages. As technology advanced, it paved the way for the development of RPA, with further notable progressions in the 1990s ("History of RPA - javatpoint," n.d.). The integration of automation to substitute human efforts is not new. It is already done at a larger scale manufacturing industry where the machines or physical robots were used instead of labor with an intention to increase the productivity, accuracy and reduce the cost. These jobs were basically blue collared jobs. However, now the new type of robots or bots are the software which are targeting the white-collar jobs by increasing knowledge intensive tasks (Chugh et al., n.d.). Here, a software is designed to take over the repetitive job and tedious job of employees which can save a lot of employees' time for more important work like analysis and decision making. Software such as workflows or bots are built for this purpose. A software developer develops a list of actions which automates the task at the backend by using the application programming interface (API). This RPA system is trained to copy the human action for a particular task in the application's graphical user interface (GUI) and train itself to repeat it directly on the GUI. The

use of automation can make it easier to integrate with the products which lack the dedicated APIs which removes the obstacles preventing such integration. These RPA helps in increasing productivity, lowering the process cost, reduction of error, improving the efficiency, and customer satisfaction and can be put to work round the clock 24 x 7 (Chugh et al., n.d.). It is predicted the RPA industry will boom from USD 2659.13 million to USD 66079.34 million industry by the year 2034 with a Compound Annual Growth Rate (CAGR) of around 37.9% (“Global Robotic Process Automation (RPA) Market Size/Share,” n.d.). The companies who have already implemented RPA have experienced the collateral benefits of standardizing and improving their processes. The basis requirement of having the structured standardized documents helps the process to be stable and helps in integration of the system. Therefore, it is reasonable to consider standardization as a benefit that arises from the intention to adopt RPA. The following characteristics are commonly found in processes that are highly compatible with RPA implementation:

- Have well-defined rules with minimal or no need for assessment.
- Contain a significant amount of manual work involved in repetitive steps.
- The processes exhibit standardization from the input stage through the process steps until the output is generated.
- Have input data available in an electronic format and not on paper.
- The process demonstrates sufficiently high transaction volumes that warrant automation.

Basically, RPA comes under Business Process Automation along with having multiple differences between the different solutions within its category. During the implementation of RPA, it is not necessary to completely modify the software codes. Instead, it can be implemented gradually or in specific sections without the need to restructure the entire process system, thereby minimizing the risk of significant business disruptions. Simple rule-based automation, advanced cognitive intelligence, and artificial intelligence are all on the automation spectrum. In case of accounting stream, it is proved that RPA has increased the efficiency as compared to human skills and reduced the number of errors of an organization. The return of investment for implementing RPA is estimated to be less than a year which depends on the complexity and the process (Wojciechowska-Filipek, 2019).

This study aims to explore the possibilities and obstacles of Robotic Process Automation (RPA) in the field of accounting, considering the time saving useful for important tasks. The primary objective is to investigate and discuss the general potential and challenges of RPA, with a particular focus on its application in typical accounting scenarios with Accounting task which is a repetitive job.

The following Thesis report includes sections like Research Question, Literature review, Research methodology, Design specification, Implementation, Evaluation and Conclusion. I will articulate the detailed study and analysis conducted throughout this academic research report.

Research Question

The research problem addressed in this study is How does the implementation of Robotic Process Automation (RPA) impact the efficiency, accuracy, and overall performance of accounting processes?

Related Work

As per the Institute for Robotic Process Automation (IRPA) definition, RPA refers to the utilization of technology that enables employees to configure computer software or robots. These configurations empower them to capture and interpret data from existing applications, perform transaction processing, manipulate data, initiate responses, and interact with other digital systems (Chugh et al., n.d.).

The GBS model disagrees from the typical approach to shared services and outsourcing services, as outlined. GBS appears when diverse processes and functions are shared within a global entity that operates functions together in a harmonious way, aiming to achieve scalability, cost reduction, and process optimization. In contrast to the traditional method of maintaining separate shared service centers and managing outsourcing vendors independently, GBS integrates business practices and locations for all shared services and outsourcing activities across organizations. It represents a synthesis of services derived from a competitive strategy that connects information and communication technology, accounting, finance, human resources, and engineering design services (Fernandez and Aman, 2018).

This article explores the Global Accounting Services (GAS) firm's provision of financial services like cash flow, investment, and decision-making to companies. Operating internationally under GBS principles, Global Accounting Services follows standardized processes and technologies. Global Accounting Services seeks cost savings, innovation, and market growth, handling complex transactions. They rely on specialized employees, their absence poses organizational risk. To enhance performance, Global Accounting Services organizations adopt disruptive technologies like robotics for accelerated processes, reflecting the growing interest in automation-driven transformations (Fernandez and Aman, 2018).

OpusCapita, a Finnish company specializing in solutions for digitalization, has demonstrated a historical trajectory that spans from its integration of electronic invoicing in the 1990s to a strategic shift towards process automation in the 2000s. Their approach to introducing Robotic Process Automation (RPA) unfolds through a structured process. Initially, a concise two-hour workshop serves as the starting point, followed by a more extensive second phase lasting approximately a day. During this phase, the focus is on meticulously evaluating existing processes by closely observing employees as they engage in task execution. Once this evaluation is complete, OpusCapita engages clients with a compelling business case, highlighting the tangible benefits of heightened productivity and reduced costs that RPA offers. Upon receiving approval for the proposed business case, the subsequent steps encompass configuring the robotic automation and ultimately delivering the solution to the client. The practical application of OpusCapita's RPA capabilities is exemplified through an illustrative portrayal found, where the implementation is bolstered by the support of UiPath. Notably, OpusCapita asserts that RPA adoption is better suited for smaller enterprises with labor-intensive routines rather than larger multinational corporations, especially those with limited tasks amenable to automation. This perspective underscores their nuanced understanding of the varying organizational contexts that determine the feasibility and benefits of RPA integration (Osman, n.d.).

O2, known as Telefonica O2, is a global telecommunications firm under the Telefonica umbrella, is one of the major players in the industry. They offer mobile plans, data services, and internet connectivity, O2 serves both consumers and businesses across several countries. Its innovative services have earned it a reputation within the telecommunications sector. Automation emerged as the next significant transformation tool for Telefonica O2. By April 2015, the company had incorporated more than 160 "robots," denoting RPA software licenses. These robots effectively managed 400,000 to 500,000 transactions every month, resulting in an impressive return on investment of 650 to 800 percent over three years. The turnaround time experienced an incredible change from days to few minutes. So, the volume of customer "chase up" calls decreased by more than 80 percent annually, as fewer clients contacted to enquire about service request status. With the robotic workforce there was scalability advantage, which was very evident, which instantly doubled the size to accommodate the launch of new products, and then effortlessly scale back after the surge subsided. This automation-driven enhancement significantly supported Telefonica O2's operational efficiency and customer satisfaction levels (Willcocks, 2015).

Methodology

Robotic Process Automation (RPA) includes a various set of methodologies that facilitate its successful implementation and execution. These methodologies present various approaches for organizations to effectively connect RPA's potential in automating their business processes. Some of the notable RPA methodologies include agile RPA, waterfall RPA, lean six sigma RPA, center of excellence RPA, proof of concept RPA, incremental and RPA. Agile RPA emphasizes adaptability and collaboration, allowing for iterative development and quick adjustments to processes as needed. Conversely, Waterfall RPA follows a sequential structure, with each phase completed before moving to the next, ensuring a well-defined plan from the outset. Lean Six Sigma RPA combines Lean principles for waste reduction with Six Sigma's focus on minimizing defects, leading to streamlined processes and high-quality outcomes. A Center of Excellence RPA establishes a hub of expertise and best practices to guide RPA implementation and scalability across an organization. A Proof-of-Concept RPA involves developing a small-scale prototype to demonstrate the viability and benefits of RPA before full-scale implementation. Incremental RPA implements automation in stages, gradually optimizing processes. These methodologies offer organizations a toolkit to tailor RPA implementation according to their unique needs, fostering efficiency, innovation, and continuous improvement. In this research, the researcher has taken the Proof-of-Concept RPA approach for understanding accounting work. This approach is carried out to get the clear picture based on human efforts and automation. This approach involves a small sample of invoices which are created by researcher in excel sheet and then converted into pdf format and then processed same data in 3 different methods.

Categories

There are three types of RPA, namely: Attended Automation, Unattended Automation, and Hybrid RPA.

Attended Automation - Attended automation involves bots which are built on the user's computer, typically initiated by the user. It proves most effective for tasks triggered at points that are challenging to detect programmatically. For instance, consider a scenario where a customer service representative must navigate through multiple screens and complete several manual steps to finalize a transaction. Rather than going through these steps, RPA robots can act as representatives, performing the required operations and seeking advice from the representative when necessary. Attended automation serves as a valuable tool to enhance tasks for customer-facing employees who still need to handle manual work (“What are the three types of RPA? - BigStep Technologies,” n.d.).

Unattended Automation - These bots operate similarly to cloud-based batch processes, seamlessly completing data processing tasks in the background. Unattended automation is a valuable solution for reducing the workload of back-office employees. It can be triggered through various methods, such as data entry at a designated location, bot startup, orchestrator startup, and at specified intervals (“What are the three types of RPA? - BigStep Technologies,” n.d.).

Hybrid RPA - In this form of RPA, a combination of attended and unattended RPA bots is used to achieve automation in both back-office and front-office tasks, facilitating end-to-end process automation (“What are the three types of RPA? - BigStep Technologies,” n.d.).

Type of Tools of Robotic Process Automation (RPA)

When seeking automation tools for your project, it's crucial to understand the different types available. Here are three categories of RPA tools:

1. **Macros** - These tools automate processes within a single system and are the most established in automation. If you need to automate a relatively straightforward process that involves only one

tool, like a spreadsheet, macros are the ideal option. Macros can efficiently organize data, perform calculations, and more.

2. IT Process Automation (IT PA) - These tools can automate intricate processes that rely on multiple systems; they fall under the category of functionally specific automation tools. IT PAs operate as behind-the-scenes bots but necessitate development and supervision from experienced IT professionals. They efficiently manage complex tasks, including handling alerts from various sources, prioritizing action plans, and executing one of your predefined possible action plans. Additionally, they promptly alert key team members who need to be informed of activities before the ticket is closed.
3. Cognitive Automation tools - These robots possess screen scraping capabilities, enabling them to mimic human actions on the desktop and perform tasks that involve multiple systems. While they might not be as precise as ITPA tools, they are user-friendly. With a straightforward drag-and-drop programming interface, they are enhanced with artificial intelligence capabilities, allowing them to learn from past examples. They also utilize natural language processing and can be referred to as intelligent or cognitive robots (“What are the three types of RPA? - BigStep Technologies,” n.d.).

Benefits of RPA

The processes automated using RPA yield several advantages.

Creates capacity from Day one – RPA offers immediate capacity creation by reducing rework time for inaccurately completed tasks and minimizing disruptions to team members, as opposed to the intricate change management inherent in traditional business process automation. The streamlined nature of RPA enables swift adoption and implementation, resulting in enhanced efficiency and improved productivity from the outset. This capacity boost leads to more resources being available for higher-value tasks and strategic initiatives, propelling organizations towards greater operational effectiveness and a competitive advantage in the modern business landscape (“The Future of Work: How Automation and AI Will Transform the Job Market’ | 01,” n.d.).

Minimizes disruption to the rest of the team - Automating repetitive tasks through RPA minimizes disruptions to the rest of the team, granting them ample time to focus on complex and non-easily automated aspects of the process. By offloading repetitive and time-consuming activities to RPA bots, team members can divert their attention to higher-value tasks, strategic planning, and problem-solving. This not only boosts team productivity but also enhances collaboration and innovation within the organization. The streamlined workflow achieved through RPA implementation fosters a more efficient and harmonious work environment, enabling teams to achieve their goals effectively while capitalizing on the advantages of intelligent automation (“The Future of Work: How Automation and AI Will Transform the Job Market’ | 01,” n.d.).

Workforce Reshaping - The process of workforce reshaping involves distribution of work within the full-time employees from routine and manual duties towards tasks that hold greater value and strategic significance. With the integration of automation, organizations free up their workforce to concentrate on activities demanding creativity, critical analysis, and innovative solutions. This approach effectively increases both overall productivity and employee contentment. This transformation optimizes the utilization of human potential and allows businesses to leverage their talent pool effectively. As repetitive tasks are automated, employees can engage in higher-level activities that contribute significantly to the organization's growth, innovation, and success in an increasingly competitive landscape (“The Future of Work: How Automation and AI Will Transform the Job Market’ | 01,” n.d.).

Increased Benefits with RPA Expansion - With the expansion of RPA implementations within the organization, a host of increased benefits emerge, particularly through the integration of unrelated programs. As RPA's scope widens, it enables seamless collaboration and synergy among different departments and processes, promoting efficiency and enhancing overall performance. The integration of diverse systems and operations streamlines data flow and communication, eliminating storage tower and improving cross-functional decision-making. Furthermore, RPA's scalability allows organizations to capitalize on its potential across various business areas, generating a cumulative effect of improved productivity, reduced costs, and heightened customer satisfaction. This progressive growth with RPA strengthens the organization's competitive edge and positions it for sustained success in the digital age.

Reduced Ongoing Upkeep - RPA offers the advantage of reduced ongoing upkeep when contrasted with conventional technologies. Due to its design as an intelligent and self-sufficient automation solution, RPA minimizes the need for constant manual intervention and updates. Unlike traditional systems, which often demand regular maintenance and updates to function optimally, RPA's adaptability and ability to learn from previous examples streamline its operation over time. This translates to lower maintenance costs, decreased downtime, and increased operational efficiency. As a result, organizations can allocate resources more strategically, focusing on innovation and growth, while RPA continues to perform reliably and consistently with minimal ongoing attention.

Self-Generated Audit Trail - RPA creates a self-generated audit trail, simplifying regulatory compliance procedures and reducing the need for additional work. By automatically recording each step of the automation process, RPA ensures complete transparency and traceability, facilitating easy auditing and compliance verification. This built-in audit trail eliminates the need for manual documentation and record-keeping, saving time and effort for compliance teams. Furthermore, it enhances accountability and risk management by providing a comprehensive history of actions taken by the RPA system. Overall, RPA's ability to generate its own audit trail empowers organizations to meet regulatory requirements efficiently while maintaining a robust and compliant operational environment. (Jovanović et al., 2019)

Reduced cost - Robotic process automation (RPA) delivers considerable advantages to companies, notably through significant cost reductions of approximately 30%. By automating repetitive tasks and streamlining workflows, RPA optimizes operational efficiency and resource allocation. The automation of manual processes translates to reduced labor costs, increased productivity, and minimized errors, resulting in overall cost savings. Additionally, RPA's ability to work 24/7 without downtime or breaks enhances productivity and accelerates task completion, further contributing to cost-efficiency. As businesses strive to optimize their operations and remain competitive, RPA emerges as a powerful solution for achieving substantial cost savings and fostering financial sustainability.

Enhance customer experience – Robotic process automation (RPA) goes beyond cost savings and plays a crucial role in enhancing the overall customer experience. By automating repetitive back-office tasks, RPA liberates valuable resources, allowing them to be redirected towards front-line customer service. This shift ensures that customers receive prompt and personalized attention, leading to improved service quality, quicker response times, and higher customer satisfaction levels. RPA's efficiency and accuracy in handling back-end operations contribute to smoother interactions and fewer errors, thereby fostering a positive and seamless customer journey. Consequently, organizations can build stronger customer relationships, bolster brand loyalty, and gain a competitive edge in the marketplace.

Lower operational risk - Robotic process automation (RPA) presents a significant advantage by eliminating operational risk, ensuring processes are executed accurately and with expertise. RPA bots

follow predefined rules and guidelines meticulously, minimizing the potential for errors and deviations that could lead to costly consequences. By automating tasks, RPA mitigates the risk of human errors and ensures consistency in process execution. Additionally, RPA maintains an audit trail, providing transparency and traceability to assess any issues promptly. This risk reduction fosters a more secure and compliant operational environment, safeguarding organizations from financial losses, reputational damage, and compliance breaches. Consequently, RPA enhances overall operational stability and resilience.

Improved internal processes - Robotic process automation (RPA) significantly enhances internal processes, with procedure definition playing a vital role. RPA implementation involves defining processes to be automated, promoting enhanced process governance and standardization. By precisely documenting each step and workflow, RPA ensures consistency, efficiency, and accuracy throughout the organization. This well-defined approach optimizes process management, reduces variability, and minimizes the potential for errors. RPA's emphasis on process governance leads to improved resource allocation, reduced operational costs, and streamlined workflows. This results in higher level of productivity and quality while adopting a culture of continuous improvement and innovation in an organization.

Integration with IT systems – One of the most important advantages of Robotic process automation (RPA) is its easy integration with existing IT systems, eliminating the necessity for replacements. RPA bots work harmoniously within the current IT infrastructure, mirroring human actions and interactions. This compatibility ensures a smooth transition to automation without disrupting established workflows or requiring extensive system modifications. RPA's adaptability extends to a wide range of applications and platforms, enabling organizations to capitalize on their existing IT investments. By leveraging RPA's compatibility, businesses can swiftly implement automation, optimize resource utilization, and achieve operational efficiency without the need for complex system repairs (“Top-5 Benefits of Robotics Process Automation (RPA) Adoption for Your Company,” n.d.).

Research method

There researcher has used Python to write the codes and complete the test of exporting the invoice details into an excel sheet using libraries like os, pdfplumber, pandas and re.

Library os is Miscellaneous operating system interfaces. This helps to use the dependent functionality like reading and writing or manipulate the file then one has to refer os.path module and in case the requirement is to read all the lines in the file then refer the fileinput module. For temporary file and directory one can refer tempfile module and for high level file and directory one can refer shutil model.

Here are the rephrased notes on the availability of these functions:

The built-in operating system dependent modules in Python are designed to maintain a constant interface across platforms. Whenever possible, they use the same interface to offer similar functionality. For instance, the function "os.stat(path)" returns stat information about the specified path in a format that originates from the POSIX interface.

Although the "os" module provides access to OS-specific extensions, using these extensions can compromise the portability of code.

All functions that accept path or file names can handle both bytes and string objects. This gives out an object of the same type as the input path or file name.

But, it's essential to note that on VxWorks, certain functions like "os.popen," "os.fork," "os.execv," and "os.spawnp" are not supported.

Similarly, on WebAssembly platforms such as wasm32-emscripTEN and wasm32-wasi, various parts of the "os" module may not be available or may behave differently. For instance, APIs related to processes (e.g., fork(), execve()), signals (e.g., kill(), wait()), and resources (e.g., nice()) may not be available. Additionally, some functions like "getuid()" and "getpid()" are either emulated or provided as stubs ("os — Miscellaneous operating system interfaces — Python 3.11.4 documentation," n.d.).

Another library that will be discussed is pdfplumber, pdfplumber is python library which is being used to extract details from pdf like text, tables, images in a programmatic way. It functions as follows:

Text extraction – pdfplumber helps extracting text from the pdf document and helps the tasks like parsing and analyzing pdf content.

Table extraction – pdfplumber is also used to extract the tables from pdf document into excel or csv format.

Image extraction – pdfplumber is used to extract images from the pdf document which helps to work on the image as a pdf content.

Page manipulation – pdfplumber can be used to work on a specific page of the pdf document and doesn't need to change everything.

Text analysis – pdfplumber helps in text analysis after extracting the text like searching for a particular word or pattern.

Pdfplumber can be used for conveniently retrieving comprehensive details about individual PDF objects. It can also be used for advanced, customizable techniques for extracting text and tables and seamless integration of visual debugging tools. Additionally practical utility functions, including object filtering using crop-box parameters.

Pdfplumber doesn't help in pdf generation, pdf modification, optical character recognition and strong support for extracting tables.

Next library in discussion is pandas

Pandas is a Python library which offers high-performance, user-friendly data structures, and tools helping in data analysis. It is an open-source and distributed under the BSD license, making it accessible for various data manipulation and analysis jobs. Pandas is mostly used for tabular data like spreadsheets, offering exploration, cleaning, and processing. Data tables are data frames, supporting various file formats for easy importing and exporting. It is an open-sourced library that introduces Dataframe and Series ("pandas documentation — pandas 2.0.3 documentation," n.d.).

DataFrame – The primary pandas data structure is a two-dimensional, size-mutable, and potentially heterogeneous tabular data. It includes labelled rows and columns, supporting arithmetic operations based on labels and can be likened to a dictionary of Series objects. It can have different variety of data types like strings, integers, and dates. Here the data can be filtered, grouped or aggregated as well ("pandas.DataFrame — pandas 2.0.3 documentation," n.d.).

Series - A Series in pandas is a one-dimensional ndarray with labelled axes, including time series. Labels can be non-unique but must be hash able. Series supports both integer and label-based indexing and provides various methods for index-based operations. When performing mathematical operations between Series, values are aligned based on their associated indexes, and the result index is the sorted union of the two indexes. Missing data is represented as NaN ("pandas.Series — pandas 2.0.3 documentation," n.d.).

Next library in discussion is re (Regular expression) operation

The re library in python is one of the powerful and flexible pattern matching tool which is used to text search, text processing and string manipulation. The re library provide various function like re.search(), re.match(), re.findall(), re.sub(), re.split() which are used for searching a pattern in a string, determining a particular pattern is matching at the beginning of the string, find all the occurrences of a pattern in a string and return as a list, replace occurrences of a pattern with a specified string, and by a specific pattern the string will be split respectively. Regular expressions use the backslash character '\ for special forms or to include special characters without invoking their meaning. In Python, this collides with the backslash's purpose in string literals. To match a literal backslash, '\\' must be used in the pattern string, as each backslash needs to be expressed as '\ inside a Python string. Additionally, note that any invalid escape sequences in Python's usage of the backslash in string literals will now generate a DeprecationWarning and may become a SyntaxError in the future. This applies even if the escape sequence is valid for a regular expression (“re — Regular expression operations — Python 3.11.4 documentation,” n.d.)

Another method is Microsoft power automate method where a process flow is created and accordingly the data is extracted from the invoices.

Design and Implementation

The objective of this research is to evaluate the impact of implementing RPA for invoice booking in the accounting software environment. The study aims to quantify the time savings, reduction in errors, and the overall improvement in accounting process efficiency achieved through RPA adoption.

In this paper, the researcher has worked on automating the invoice booking into the Bookkeeping software “Xero” which saves ample time of an accountant doing a repetitive job which can be used for doing important and brain storming activities. In this paper, invoices are processed using multiple methods i.e., Microsoft Power Automate, Python and Manual booking of invoices and has done a comparable study and came up with the inference under which circumstance which methods are useful and benefits of same.

The manual invoice booking process requires entering invoice details into the accounting software Xero manually. It relies on human involvement and can be a time-consuming task. The individual responsible for this activity must meticulously input all relevant information, such as invoice number, date, vendor details, item description, and amounts. While manual invoice booking provides a heightened level of control and precision, it is vulnerable to errors that may occur due to human errors or oversight.

Manual invoice booking in Xero accounting software involves the process of entering and recording bills manually that are not automatically imported into the system. This can be done in a few simple steps.

Step 1 - Firstly, log in to the Xero accounting software using unique credentials. Once logged in, navigate to the "Dashboard" tab, depending on the layout of the software.

Step 2 - Within the "Business" tab, there will be option to click Bills to pay. Clicking on this button will begin the process giving two option New Bill and Import bill. Alternatively, some versions of Xero may have a quick add button directly on the dashboard.

Step 3 - A new screen will appear, prompting you to fill in the necessary details of the invoice. Starting by selecting the vendor name from whom the invoice is being received. If the Vendor does not exist in system, one may need to add their details as a new contact.

Step 4 - Next, the invoice number, date, and reference as stated in the invoice from the vendor needs to be entered. It needs to be ensured that the invoice number is added correctly for accurate record-keeping.

Step 5- Moving on, it is crucial to input the individual line items or products/services are also recorded. For each line item, a description should be provided, quantity, rate, account number and relevant tax information if applicable.

Step 6 - Discounts, quantities, or line items can be removed as per the requirement. Below the line items, there is usually a section to add additional notes or terms specific to the invoice or client. This is optional but can provide further clarity if required in the future.

Step 7 - After entering all the required information, it needs to be carefully assessed that the invoice accurately entered and completed. Thorough verification of the calculations and prompt rectifications of required needs to be taken.

Step 8 - After this verification of details there will be options like "Save", "Approve" and "Cancel" button to finalize the process.

Step 9 - Saving the invoice means it will be stored in the system for future reference and needs to be approved to get to loaded into the accounting ledger whereas if "Approve" is clicked the entry will be posted into the ledger and can be viewed in Trail Balance.

By following these steps, anyone can manually book invoices in Xero accounting software. This process ensures that all financial transactions are accurately recorded, providing a clear audit trail and comprehensive financial records.

A link of video presentation of same is pasted here for reference.

[Manual invoice booking in Xero.mp4](#)

Invoice booking using Microsoft Power Automate introduces advanced automation capabilities, revolutionizing the traditional process. The system leverages predefined templates and workflows to swiftly extract vital details from invoices, encompassing vendor information, line items, and respective amounts. This automated approach significantly diminishes the reliance on manual data entry, substantially expediting the entire invoice booking procedure. By reducing human intervention, Power Automate enhances efficiency, minimizing the risk of errors and delays inherent in manual workflows. Despite its automation prowess, the process does not entirely remove the role of human involvement. Human verification and oversight remain integral to guarantee the accuracy of extracted data and to address any potential discrepancies or exceptions. Even with accurate templates and automated data extraction, certain invoices may pose unique challenges, necessitating human expertise to ensure precise and reliable results.

Incorporating both automation and human intervention, this hybrid approach harnesses the strengths of each element. The automated aspect optimizes speed and consistency, while human oversight provides a critical layer of scrutiny and validation. Collaboratively, they create a harmonious and robust system that not only accelerates the invoice booking process but also ensures the reliability and integrity of the financial data.

By embracing Microsoft Power Automate for invoice booking, organizations unlock the potential for streamlined operations, reduced workload, and enhanced accuracy in financial processes. This combination of technological prowess and human expertise signals a new era in accounting and finance, where efficiency and precision walk hand in hand.

Instead of manually loading the invoices in Xero, we have to click on the option of Import bill which will give the template that is acceptable by Xero, and we have to extract the data from the Vendor invoices into this form and import it into the system.

Using Microsoft Power Automate to book a vendor bill in Xero accounting software involves creating a flow that automates the process of entering the bill details. Following are the details listing the step-by-step process to set up this automation flow:

Step 1: Login to Microsoft Power Automate – Start the Microsoft Power Automate website (flow.microsoft.com) and log in to Microsoft account or work/school account.

Step 2: Create a New Flow - Click on "All Apps" in the top navigation bar, then select "Power Automate" and Explore under AI Builder to create a new flow from scratch.

Step 3: Set the Trigger – Select Extract information from Invoices. Upload an invoice and select Use prebuilt model, Use in a flow and create a flow and the format in excel sheet.

Step 4: Add Action – Specify the details from PDF invoices to be extracted into excel format and train the flow.

Step 5: Map the Fields - Power Automate will ask you to provide the necessary information for extracting the bill data. Map the fields from your trigger (e.g., form fields or system data) to the corresponding fields in Xero via excel sheet.

Step 6: Save and Test the Flow - Once the flow is completed the configuration, click on "Save" to save the flow. Now the flow can be tested by triggering it with a test input (e.g., simulate form data) or wait for the actual trigger event to occur.

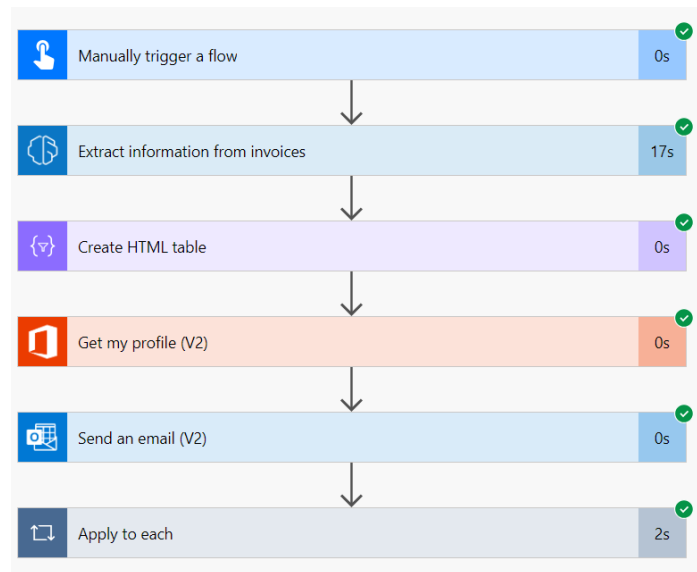
Step 7: Review and Confirm - Review the bill data extracted into excel sheet and ensure all the information is accurate. If needed, adjustments to the flow settings or mapping can be done.

Step 8: Turn on the Flow - Once it is satisfied with the setup and testing, turn on the flow. From now on, whenever the trigger event occurs or when we trigger it manually (e.g., new form submission), Power Automate will automatically extract the Vendor bill data from PDF to excel sheet which can be imported in Xero based on the provided information.

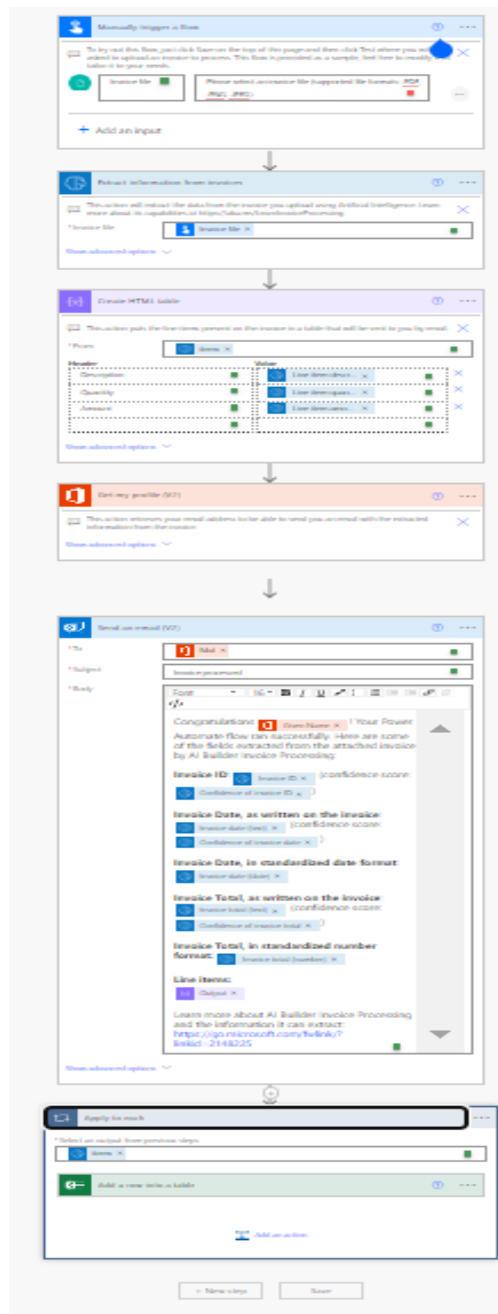
A link of video presentation of same is pasted here for reference.

[MS Power Automate flow and extracting invoices details.mp4](#)

Screenshot of the flow



A detailed screenshot of the above flow explaining each step



Invoice booking using Python into Xero accounting software takes automation a step further. Python, a powerful programming language, can be used to develop custom scripts or programs that can extract data from invoices and export it into excel sheet which can be imported in Xero. Python scripts can be designed to accurately extract data from invoices, match them with existing vendor records in Xero, and post the information. This method not only saves time but also minimizes errors by ensuring consistent and accurate data transfer.

Invoice booking in the accounting software Xero can be accomplished using Python code in a step-by-step manner. Presented below is a comprehensive guide detailing the process to achieve this.

1. Install the Python packages:

To start with researcher has used Google colab for Python and installed the required Python packages, such as pandas, os, re and pdfplumber which allows the data reading, manipulation and data extraction.

2. Load the invoices into the directory

Create a folder and load all the invoices to be processed.

3. Process the codes

Once all the invoices are loaded in the directory, process the codes by giving the path of the folder and as an output an excel will be generated which is a summary of invoices that is loaded.

4. Retrieve and reformat the details:

Similar to Power Automate method, the extracted details need to be reformatted, click on the option of Import bill which will give the template that is acceptable by Xero, and we have to extract the data from the Vendor invoices into this form and import it into the system.

5. Verify

Once the invoices are imported as per the specified template all the invoices will be imported and one can verify the details by matching the balance of vendor or by Bank reconciliation.

Automation experts predict a rising economic significance of RPA in the years ahead. They foresee its accessibility expanding to various industries, leading to substantial growth in the RPA market during the 2020s. Unlike conventional automation, RPA software possesses the ability to learn from errors, handle unexpected issues, and adjust to rapid business model changes. As continuous improvements occur, RPA and similar intelligent technologies will become essential tools for businesses of all sizes. Industries that fail to leverage RPA might face a significant disadvantage compared to those embracing modern computerized automation (“Robotic process automation (RPA) . - EBSCO,” n.d.).

When comparing these three approaches, it becomes evident that manual invoice booking is the most time-consuming and prone to errors. While Microsoft Power Automate offers automation capabilities, its reliance on predefined templates and the need for human verification can still result in occasional errors. On the other hand, invoice booking using Python provides the highest level of automation, accuracy, and efficiency. By leveraging Python's programming capabilities, custom scripts can directly integrate invoice data into Xero, eliminating the need for manual input or human validation.

In conclusion, while each of these methods has its advantages and disadvantages, invoice booking using Python into Xero stands out as the most efficient and accurate approach. By automating the entire process, Python minimizes human errors and ensures the seamless integration of invoice data into Xero, ultimately enhancing the overall efficiency of the accounting process.

Implementation guidelines for RPA and future outlook

Automation acquisition of data from paper documents

In this world of business there are many businesses which run on hand made invoices and such documents are scanned and converted to pdf manually. This manual data entry process incurs considerable costs without yielding any additional benefits. To address this inefficiency, a Power Automate-based approach can be utilized for automating the extraction of invoice details from PDF documents and seamlessly populating them into an Excel spreadsheet.

By leveraging Power Automate's capabilities, the process of invoice details extraction becomes more efficient, significantly reducing the need for manual data entry in public administration workflows. This approach streamlines the extraction of relevant data such as invoice numbers, dates, vendor details, item

descriptions, and amounts from PDF documents. Consequently, the automation contributes to substantial cost savings and improved productivity.

However, it is essential to acknowledge that automated data extraction from PDFs may still result in occasional inaccuracies, depending on the quality and structure of the PDF documents. For instance, handwritten or poorly formatted data in the PDF may pose a higher probability of errors during automated extraction. To ensure data accuracy, human oversight remains necessary to validate the extracted invoice details and address any discrepancies or exceptions.

Incorporating Power Automate for invoice details extraction from PDF to Excel brings efficiency and accuracy to recording the accounts in the business and administering the financials. This combination of automation and human verification creates a harmonious workflow, optimizing productivity and maintaining the integrity of invoicing data in the context of business administration.

Automated integration of data

From an accounting perspective, employing Process Automation (PA) methods can facilitate the mixture of relevant data from various systems into a centralized and process-oriented data repository. This repository contains all necessary information required for the smooth execution of accounting and financial tasks. In such a setup, Robotic Process Automation (RPA) systems can play a crucial role in automatically ensuring data consistency across all interconnected and associated databases. This application of RPA presents a significant opportunity for enhancing operational efficiency.

Automated integration of processes

RPA offers valuable support in streamlining administrative processes by automating the initiation of subsequent tasks or directing tasks to the relevant unit responsible for further processing, thereby minimizing waiting times. These scenarios are often implemented through integrated workflow management systems. However, within the accounting stream, the situation can be more complex, with different application systems handling various aspects of a work process or even some parts lacking support from any information system. In such instances, RPA becomes instrumental in achieving an "integrated" automated workflow management, despite the absence of a unified information system. By leveraging RPA, organizations can bridge the gap between unrelated systems and create a cohesive, automated workflow that ensures seamless task handoffs and efficient process execution. This approach optimizes resource utilization, reduces delays, and enhances overall productivity, even in situations where complete integration across information systems may not be feasible like accounting and auditing.

The flexible nature of RPA allows it to adapt to diverse process landscapes and establish efficient connections between different components, enabling accounting forms and companies to enhance their operational efficiency and service delivery without overhauling their existing systems.

Evaluation

The researcher established a test company called "Research Company" within the Xero platform, specializing in the Fruit Supply industry. To facilitate testing, a set of 108 test invoices were created, following a standardized format. The invoice template encompasses several key criteria, including Contact Name, Invoice Number, Invoice Date, Due Date, Description, Quantity, Unit Price, Total Amount, Gross Amount, and Tax. These criteria played a crucial role in shaping the flow and development of the Python code to handle the invoicing process effectively. By adhering to this comprehensive template, the researcher ensured consistency and accuracy in generating the test invoices, making it easier to assess the functionality and performance of the system.

The researcher has generated a series of 108 invoices, sequentially numbered from 1 to 108. Initially, the first 12 invoices were manually entered into the Xero system. The time taken to record these 12 invoices was used as a baseline for calculating the overall time required to process all 108 invoices. Subsequently, the next set of 36 invoices was utilized to test the Power Automate flow. Again, the time taken for recording this subset of invoices served as the reference for estimating the time needed to process all 108 invoices using the Power Automate flow.

Finally, the researcher applied the Python code to test all 108 invoices. The time taken for this complete invoicing process using Python was also evaluated. By comparing the timings from each stage, the researcher could draw inferences about the efficiency and effectiveness of each approach in handling the invoicing tasks. This comprehensive evaluation allowed for a thorough assessment of the manual entry, Power Automate, and Python-based invoicing methods.

Invoice booking – Manual invoice

The researcher manually booked 12 invoices in Xero Accounting software within 13 minutes. On scaling, it would take approximately 1 hour and 57 minutes for the same person to complete the task of booking all 108 invoices.

Invoice booking – Power Automate

The researcher created a workflow by using the Power Automate App to input and process 36 invoices, subsequently extracting a summary from the Excel sheet. The processing time for each individual invoice averaged around 18 seconds, resulting in a cumulative time of approximately 12 minutes for handling the batch of 36 invoices. Scaling this up to 108 invoices, the estimated processing time would extend to around 36 minutes. Additional processing encompassing the extraction into the Xero template and data integration into Xero itself was completed within about 10 minutes. Overall, the Power Automate method demonstrated an execution time of 46 minutes. This accomplishment signifies a notable 61% reduction in time when contrasted with the conventional manual approach to invoice processing. The implementation of the Power Automate technique not only enhances efficiency but also significantly accelerates the processing of invoices, making it a valuable advancement in workflow optimization.

The screenshot shows the configuration page for a Power Automate flow named "Read information from invoices - 4". The flow is currently "On" and was created on August 13, 12:08 PM. The description states: "Click a button to quickly extract information from invoices using Artificial Intelligence and send the results by email. First time using AI Builder Invoice Processing? Learn more at https://aka.ms/LearnInvoiceProcessing". The owner is Veena Hanumanthu Chenneti.

The "28-day run history" table shows the following data:

Start	Duration	Status
Aug 13, 12:20 PM (20 min ago)	00:00:18	Succeeded
Aug 13, 12:20 PM (21 min ago)	00:00:18	Succeeded
Aug 13, 12:20 PM (21 min ago)	00:00:18	Succeeded
Aug 13, 12:20 PM (21 min ago)	00:00:18	Succeeded
Aug 13, 12:19 PM (21 min ago)	00:00:18	Succeeded
Aug 13, 12:19 PM (21 min ago)	00:00:18	Succeeded

Other details visible in the screenshot include:

- Connections:** Microsoft Dataverse, Excel Online (Business), Office 365 Outlook (Permissions), and Multiple permissions.
- Process mining (preview):** Average run duration of 00:00:18.
- Run only users:** No users are currently associated with the flow.
- Associated Apps:** No apps are currently associated with the flow.

Invoice booking – Python

Using Python for invoice processing, the task of loading 108 invoices into the Google Colab directory, including folder creation, invoice loading, and code execution, was completed in approximately 2 minutes. The outcome was an Excel file featuring 108 line transactions. Subsequently, the established Excel file underwent conversion to CSV, aligned with Xero's import template, requiring around 8 minutes. Once this conversion was accomplished, the import process, spanning all 108 invoices into Xero, concluded within 13 minutes. This approach showcases the efficiency of Python-based automation in handling invoice processing tasks, streamlining various stages from initial loading to final integration with accounting software.

Conclusion and Discussion

This research work was aimed at measuring the time required to manually book the invoices into accounting software xero and three separate methods were tested i.e. manual invoice booking, extracting invoice data using power automate and then loading the file into xero and third method using python libraries extracting the invoice details into excel and then loading them in xero.

The manual invoice processing task of manually inputting 12 invoices through Xero Accounting software, achieving completion within a 13-minute timeframe. By reasoning from this experience and utilizing the foundation of 12 invoices, one could estimate that the researcher's work to book the entire 108 invoices would likely take around 1 hour and 57 minutes. The effectiveness and precision of manual invoice booking are linked to the person's competence. While the initial pace might be slow, a noticeable improvement typically surfaces as the individual becomes more familiar with the procedure.

Invoice booking using the Power Automate App, the researcher designed a system and processed 36 invoices. After inputting and processing them, a summary was taken from the Excel sheet. Each invoice took about 18 seconds on average, and in total, it took around 11 minutes and 24 seconds to deal with all 36 invoices together. If this process was done for 108 invoices, it would take about 34 minutes and 12 seconds. Then, more work was done – the information was moved from Excel to Xero. This part of the process took about 10 minutes. When everything was added up, the whole method using Power Automate took around 40 minutes. This achievement is important because it means that using Power Automate is 61% quicker compared to doing things manually, the usual way.

When the Python code was used to handle a bunch of 108 invoices, something impressive happened. In just 2 minutes, the Python code took a small piece of information from the Excel sheet, like the price and quantity of things. This was way faster and more accurate compared to doing it by hand like usual. The way Python worked was so much better than the normal way of doing things. The time it took to do all this was much shorter, only about 1/9th of the time it used to take. This shows how using Python to help with work can make things way faster and smoother.

So, in the end, using Python to get important information from invoices in just 2 minutes is a big change because of technology and the overall process is completed in 13 minutes. But, what's even more important is how well it fits with each business, like "Research," and how it can be made just right for them.

As per the researcher's evaluation, among the three methods, the Python approach using the library pdfplumber approach emerged as the most viable option. This method effectively reduces the time required compared to manual invoice booking, all while ensuring the integrity and comprehensiveness of the data, efficiency, accuracy and overall performance of accounting process in the business context.

Future work – There is scope of using VBA codes for the manual process of converting the data in excel file that is extracted into csv so that the complete process is automated and only loading the files in Xero and approving would be a manual process. Another future scope is API integration of python with Xero software for a complete change of scenarios.

Bibliography

- Chugh, R., Macht, S., Hossain, R., n.d. Robotic Process Automation: a review of organizational grey literature. *International Journal of Information Systems and Project Management* 10. <https://doi.org/10.12821/ijispm100101>
- Fernandez, D., Aman, A., 2018. Impacts of Robotic Process Automation on Global Accounting Services. *Asian Journal of Accounting and Governance* 9, 123–131. <https://doi.org/10.17576/AJAG-2018-09-11>
- Global Robotic Process Automation (RPA) Market Size/Share [WWW Document], n.d. URL <https://www.globenewswire.com/news-release/2023/03/24/2633986/0/en/Global-Robotic-Process-Automation-RPA-Market-Size-Share-Expected-to-See-USD-66-079-34-Million-by-2032-at-37-9-CAGR-Polaris-Market-Research.html> (accessed 7.16.23).
- History of RPA - javatpoint [WWW Document], n.d. URL <https://www.javatpoint.com/history-of-rpa> (accessed 7.23.23).
- Jovanović, S., Jovanović, S.Z., Đurić, J.S., Šibalija, T. V, 2019. ROBOTIC PROCESS AUTOMATION: OVERVIEW AND OPPORTUNITIES. *International Journal of Advanced Quality* 46.
- os — Miscellaneous operating system interfaces — Python 3.11.4 documentation [WWW Document], n.d. URL <https://docs.python.org/3/library/os.html> (accessed 7.29.23).
- Osman, C.-C., n.d. Robotic Process Automation: Lessons Learned from Case Studies. <https://doi.org/10.12948/issn14531305/23.4.2019.06>
- pandas documentation — pandas 2.0.3 documentation [WWW Document], n.d. URL <https://pandas.pydata.org/docs/> (accessed 7.29.23).
- pandas.DataFrame — pandas 2.0.3 documentation [WWW Document], n.d. URL <https://pandas.pydata.org/docs/reference/api/pandas.DataFrame.html> (accessed 7.29.23).
- pandas.Series — pandas 2.0.3 documentation [WWW Document], n.d. URL <https://pandas.pydata.org/docs/reference/api/pandas.Series.html> (accessed 7.29.23).
- re — Regular expression operations — Python 3.11.4 documentation [WWW Document], n.d. URL <https://docs.python.org/3/library/re.html> (accessed 7.31.23).
- Robotic process automation (RPA) . - EBSCO [WWW Document], n.d. URL <https://research.ebsco.com/c/x47ol5/viewer/html/6hznhez3yj?auth-callid=510f2022-5675-9b68-8622-ce06d51d906a> (accessed 7.26.23).
- Syed, R., Suriadi, S., Adams, M., Bandara, W., Leemans, S.J.J., Ouyang, C., ter Hofstede, A.H.M., van de Weerd, I., Wynn, M.T., Reijers, H.A., 2020. Robotic Process Automation: Contemporary themes and challenges. *Comput Ind* 115, 103162. <https://doi.org/10.1016/J.COMPIND.2019.103162>

“The Future of Work: How Automation and AI Will Transform the Job Market” | 01 [WWW Document], n.d. URL <https://vocal.media/01/the-future-of-work-how-automation-and-ai-will-transform-the-job-market-fq2mp009y> (accessed 7.27.23).

Top-5 Benefits of Robotics Process Automation (RPA) Adoption for Your Company [WWW Document], n.d. URL <https://www.provintl.com/blog/top-5-benefits-of-robotics-process-automation-rpa-software> (accessed 7.26.23).

What are the three types of RPA? - BigStep Technologies [WWW Document], n.d. URL <https://bigsteptech.com/blog/types-of-rpa/> (accessed 7.26.23).

Willcocks, L., 2015. The Outsourcing Unit Working Research Paper Series Paper 15/07 Robotic Process Automation: The Next Transformation Lever for Shared Services.

Wojciechowska-Filipek, S., 2019. “Automation of the process of handling enquiries concerning information constituting a bank secret” NUMBER OF REFERENCES 28 NUMBER OF FIGURES 3 NUMBER OF TABLES 0 Automation of the process of handling enquiries concerning information constituting a bank secret. Banks and Bank Systems 14. [https://doi.org/10.21511/bbs.14\(3\).2019.15](https://doi.org/10.21511/bbs.14(3).2019.15)

Appendix

Libraries which were tried before settling for pdfplumber and openpyxl are PyPDF2. This library provides functions and classes that enable you to extract, manipulate , and create PDF documents.

```
import openpyxl
from PyPDF2 import PdfReader

# Open the PDF file
with open('Arkway 15 june invoicetemplate (1).pdf', 'rb') as file:
    # Create a PdfReader object
    pdf_reader = PdfReader(file)

    # Extract text from the PDF
    text = ""
    for page in pdf_reader.pages:
        text += page.extract_text()

# Find the Invoice number
Invoice_number = None
lines = text.split('\n')
for line in lines:
    if 'Invoice_number' in line:
        Charges = line.split(' ')[1].strip()
        break

# Find the Invoice mo
Invoice_no = None
lines = text.split('\n')
for line in lines:
```

```

        if 'Invoice_no' in line:
            Charges = line.split(' ')[1].strip()
            break

# Find the Charges total
Charges = None
lines = text.split('\n')
for line in lines:
    if 'Charges' in line:
        Charges = line.split(' ')[1].strip()
        break

# Find the invoice total
VAT_total = None
lines = text.split('\n')
for line in lines:
    if 'VAT' in line:
        invoice_total = line.split(' ')[1].strip()
        break

# Find the invoice total
invoice_total = None
lines = text.split('\n')
for line in lines:
    if 'Total' in line:
        invoice_total = line.split(' ')[1].strip()
        break

# Create a new Excel workbook
workbook = openpyxl.Workbook()

# Select the active sheet
sheet = workbook.active

# Write the invoice total to the Excel sheet
sheet['A1'] = 'Invoice Total'
sheet['A2'] = invoice_total
sheet['B1'] = 'Charges'
sheet['B2'] = Charges
sheet['C1'] = 'VAT'
sheet['C2'] = VAT_total

# Save the workbook
workbook.save('Arkway_total.xlsx')

```

But these codes had error which were as follows:

No table found in /content/Arkway 15 june invoicetemplate (1).pdf.
Skipping...

```
-----  
-----  
NameError                                Traceback (most recent call  
last)  
<ipython-input-53-6ab6ffe59cc3> in <cell line: 32>()  
    31 invoice_summary = {  
    32 'Invoice Number': filename,  
--> 33 'Total Amount': invoice_table['Total'].sum()  
    34 }  
    35 invoice_data.append(invoice_summary)  
  
NameError: name 'invoice_table' is not defined
```

One more library that I had worked and tried executing was tubula-py.

```
import tabula  
import pandas as pd  
  
# Path to the PDF file  
pdf_file = "/content/BalanceAndTransactionReport (17).pdf"  
  
# Specify the page number(s) containing the table(s) you want to  
extract  
pages = "1-2"  
  
# Extract the table(s) from the PDF  
df = tabula.read_pdf(pdf_file, pages=pages)
```

and the error is as follows

```
FileNotFoundError                          Traceback (most recent  
call last)  
<ipython-input-3-cc0dfd135985> in <cell line: 11>()  
    9  
   10 # Extract the table(s) from the PDF  
--> 11 df = tabula.read_pdf(pdf_file, pages=pages)  
  
/usr/local/lib/python3.10/dist-packages/tabula/io.py in  
read_pdf(input_path, output_format, encoding, java_options,  
pandas_options, multiple_tables, user_agent, use_raw_url, pages,  
guess, area, relative_area, lattice, stream, password, silent,  
columns, relative_columns, format, batch, output_path, options)  
   417  
   418     if not os.path.exists(path):  
--> 419         raise FileNotFoundError(errno.ENOENT,  
os.strerror(errno.ENOENT), path)  
   420  
   421     if os.path.getsize(path) == 0:  
  
FileNotFoundError: [Errno 2] No such file or directory:  
'/content/BalanceAndTransactionReport (17).pdf'
```


One more code using tabula

```
import tabula
import pandas as pd

# Define the path to your PDF file
pdf_file = "//content/Test Exp invoices8.pdf"

# Extract tables from PDF into a list of DataFrames
dfs = tabula.read_pdf(pdf_file, pages='all', multiple_tables=True)

# Filter out None values (pages without tables)
dfs = [df for df in dfs if df is not None]

# Create a new Excel workbook
workbook = pd.ExcelWriter("output.xlsx", engine='openpyxl')

# Write each DataFrame as a separate sheet in the Excel file
for i, df in enumerate(dfs):
    df.to_excel(workbook, sheet_name=f"Page{i+1}", index=False)

# Save the workbook
workbook.save()
```

error – the file was generated but had no data

In the scope of manual invoice booking, a set of 36 invoices were successfully processed. This process was completed without any errors, which might be because of the relatively modest number of invoices involved. Nevertheless, it remains significant to acknowledge the characteristic vulnerabilities inherent to human-driven tasks, such as the details of invoice booking. The repetitive nature of this task can delay the provisions of constant 100% level of concentration. As monotony sets in, the chances for errors arise. Consequently, relying exclusively on manual efforts for processing of invoice booking doesn't assure accuracy threshold of 100%. The complexity required in accurate invoice booking is combined with the limitations in human performance. The occurrence of careless errors within this domain, whether due to lapses in concentration or the repetitive nature of the task, can bring in financial discrepancies and operational inefficiencies. Hence, the necessity to minimize the potential for errors necessitates a fundamental shift towards more efficient and automated approaches. Incorporating automated technologies, such as digital systems and software solutions, can introduce higher levels of precision and reliability in financial processes like invoice booking. This, in turn, diminishes the potential for discrepancies, boosting the financial integrity of the organization. As the landscape of modern business continues to evolve, the integration of automated solutions emerges as a strategic imperative to enhance accuracy, reduce risks, and optimize overall operational efficiency.

Robotic Process Automation significantly boost's employee motivation within an organization. By relieving employees of repetitive and monotonous tasks, RPA allows them to redirect their energy towards more creative, strategic, and intellectual activities. This shift not only utilizes their capabilities but also empowers them to contribute meaningfully to the organization's goals. It also boosts employee morale through upskilling and sideways skill development while automating

monotonous tasks. When employees get free time from their monotonous jobs which will have less stress on them which can further help the organization with new ideas and making important decisions. This will make them feel important as well as proud of their work. Also, when they see how much better things are with RPA, like getting things done faster, making fewer mistakes, and having simpler processes, it can make them feel happier and more involved at work. This transformation encourages a renewed sense of purpose, as employees recognize that their contributions are essential to driving the successful integration and utilization of automation technologies. Ultimately, RPA empowers employees to become drivers of positive change and innovation, elevating their motivation levels and overall job satisfaction.

When new technologies are introduced, people often worry about their jobs staying secure. This is something that has happened before with every innovation and automation. Even when excel was introduced people worried about their jobs. When things like computers and machines change, it can make some jobs and money situations not so good. But this doesn't mean people will always lose their jobs. Sometimes, these changes can also make new jobs and make old jobs different. So, companies should help their workers. They should teach them how to use new things, explain the good stuff these new things can do, and tell them about the things these new things can't do. This helps employees be ready for the changes that will happen. Furthermore, as the accounting field gets simpler with the introduction of technology in it, this research found that accountants need to know about information technology along with their regular finance and accounting skills. Rather than sticking to the same tools for long, accountants need to upgrade themselves with new technology every year. This doesn't only mean knowing how the technology works, but also being able to get used to it as time goes on as it affects how successful employees and companies are.

Consequently, relying exclusively on the speed set by processing 12 invoices to predict the timeframe required for 108 invoices might not provide a comprehensive view of the scalability process. The dynamics of scaling can introduce variables that influence the actual efficiency and time taken. A more complex workload could impact the pace achieved, and familiarity with the task might not scale linearly. Additionally, external factors like exhaustion or interruptions could further change the projected timescale. Therefore, while projecting from the base of 12 invoices provides a rough estimate, it is essential to consider the extreme detailed dynamics of scaling to arrive at a more accurate prediction of the time required for processing a larger quantity of invoices.

Using Power Automate doesn't only save time, but it also makes dealing with invoices much faster and smoother. In the end, the Power Automate method is a big step forward. It's not just about making things more efficient, but also about getting invoices processed much faster. This is a big advantage for making work processes better and faster.

But it's also important to remember that this success should be thought about in terms of each business's own way of working. For example, in a business like "Research" that focuses on apples, the quantity and number of apples is important. Even though doing things in just 1/9th of the time is good, it's even more important to think about how it fits with the special needs of that business. This means that while using Python to make things faster is great, it's even better when it's adjusted to match what each business needs.