

Invoice Processing Using Robotic Process Automation

Sagar Sahu, Sania Salwekar, Atharva Pandit, Manoj Patil

Computer Engineering, Datta Meghe College of Engineering, Airoli, Navi Mumbai, Maharashtra, India

ABSTRACT

This paper describes our recent effort to develop a automative application to transform invoice processing in Finance operations. As a prime example of the technology's potential for driving efficiency, Robotic Process Automation (RPA) can be applied to a number of finance and accounting operations, invoice processing. RPA DataBot can automate data input, error reconciliation, and some of the decision-making required by finance staff when processing invoices. At the same time, automation is able to limit errors in such processes and reduce the need for manual exception handling. UiPath's RPA DataBot are able to constantly monitor a dedicated folder where invoices are saved by employees (or other DataBot) in PDF format. Once robots detect the presence of an invoice in the folder, they begin to extract information from the document. Using intelligent Optical Character Recognition (OCR) and natural language processing capabilities, DataBot are able to read out the information that is visible on the invoice. After robots extract the key information from each invoice, they use their credentials to open the company's database or enterprise resource planning system, if not already open. The robots then start processing the invoices one-by-one by transferring over the relevant invoice information. During this whole process, the DataBot are also running background activities such as monitoring the dedicated invoice folder or its email address, performing basic checks to see if the company's database is open, and verifying whether vendor information (e.g. VAT number) on the invoice matches what is already in the database.

Keywords : DataBot, Optical Character Recognition, Robotic Process Automation

I. INTRODUCTION

According to a research report by technology services company Aberdeen Group it takes between 4 to 16 days for companies to process an invoice from receipt through payment approval. This conclusion might not really come as a shock, though especially for companies and their employees who experience the manual effort still involved with processing invoices on a day-to-day basis. In fact, a study by Canon Business Process Services suggests that more than half of all invoice processing requires at least 86-100% manual input.

In times of increasing digitization and automation, these numbers reveal that companies are missing out on possibilities for making their finance & accounting

departments more efficient. And the high percentages of manual labor indicate significant inefficiencies combined with a lack of central or standardized invoice management processes, which can also lead to unnecessary inconveniences like unpaid invoices (and the associated late fees) as well as higher than necessary labor costs.

1. Automation's potential with invoice processing

Financial processes are critical to the overall functionality of companies, which means that there may be doubts regarding the capabilities and success rates of RPA for accounts payable activities like invoice processing. Invoices that company receive from their suppliers sometimes, for example, arrive in multiple different formats: as a paper copy, a Word document, a

PDF email attachment, or a fax. Moreover, a company's finance team is responsible for manually transferring the data from these various invoice formats into the company's database and dealing with any discrepancies.

These challenges make it difficult to handle invoices in the same way each time and, as a result, drive company employees to process invoices manually. Moreover, because the data in these invoices is typically unstructured, this manual transfer process is time-consuming and prone to errors. But do these operations need to be done manually? In short, the answer — perhaps unsurprisingly — is no

2. Invoice processing with UiPath

2.1 Fast, measurable outcomes :- Let's see how businesses can leverage UiPath's Enterprise RPA Platform in driving down the costs and simultaneously increasing the efficiency and accuracy of their accounts payable operations.

Global BPO provider uses UiPath for Citrix intensive automation.

A global BPO provider used UiPath to reduce the delivery effort in several of their clients' invoice process. This required access to their ERP system through remote virtual environment (citrix technology). UiPath's ability to precisely recognize screen elements and drive seamless automation even in the toughest scenarios such as Citrix helped the BPO to reduce the delivery effort by 70%, with a drop in AHT (average handling time) from 10 minutes to 3.

European Pharma company uses intelligent OCR to process invoices in unstructured form. A European Pharma company was using 3 FTEs to process 1800 invoices every month. Four applications, including MS Office and e-mail, were involved, and several document types, many of the invoices still being in paper format. The UiPath Robot extracted the structured and semi-structured data using intelligent OCR and performed all validations necessary, such as the PO number, the number of items, and the cost per item against the PO and goods receives.

Figure 1 shows global growth in revenue using Robotic Process Automation use cases.

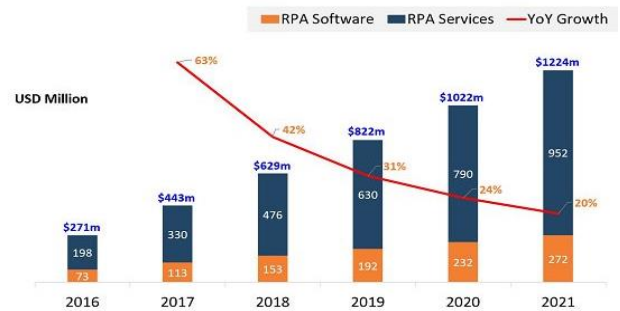


Fig 1. Global Robotic Process Automation (RPA) Market.

II. LITERATURE REVIEW

The invoicing process is an important part of a wider set of business processes including the placing and acceptance of an order, delivery and payment. Some research papers regarding invoice processing system were studied by us. These included various techniques of invoice processing using cognitive approach, generic system which uses OCR engine, using CBR (Case-based reasoning), optimization approach which makes use of SVM, Entropy, etc. Some of these approaches for invoice processing system are mentioned below.

Enes Aslan, Ethem Unver, Tugrul Karakaya, Yusuf Sinan Akgul in their paper proposed a new invoice parsing method which consists of two-phase optimization structure and eliminates invoice classes. The first phase uses individual invoice part detectors such as SVM, maximum entropy and HOG to produce candidates for the various parts of different types of invoices. At the second phase, the basic idea is to divide an invoice into different parts and then arrange it together. As PBM is an optimization based method, it can handle any type of invoices. The proposed system is tested with real invoices and are found to be promising for the real-world experiences.

Ying Li, Muthu Muthiah, Arindam Routh, Chitra Dorai describes their initiative of developing a cognitive application to transform invoice processing system in P2P (Procure to Pay) process by providing domain-

specific systems. Procure to pay (P2P) process consists of two main functions i.e. transactional data is sent to supplier as well as it includes handling of data which includes information about orders and payment for a product or service. A 3-month long usage tracking has showed promising performance. Significant productivity improvement has been achieved by the team since the deployment of this application in production with a client.

T. A. Bayer, H. U. Mogg-Schneider, Daimler-Benz proposed a generic system for processing invoices that automatically extracted the requested items with arbitrary form layout in arbitrary domains. The given system consists of two components OCR tool and FRESCO. The error rate is less for FRESCO so it is used. FRESCO component is contains the knowledge about the domain. This generic system is used for health insurance related invoices.

Hatem Hamza, Yolande Belaid, Abdel Belaid introduced an invoice analysis approach using Case-Based Reasoning (CBR). CBR is used to analyse and interpret new invoices using previous processing experiences. Each new documents are segmented into structures and interpreted using structure database. A new document's interpretation depends on graph edit distance as well as on string edit distance. They focus more on document interpretation through its structure interpretation. Mainly, CBR approach was used for accessing multiple invoices. The final use of this system was processing invoices for both known and unknown classes.

III. REQUIREMENT ANALYSIS

Requirements analysis involves all the tasks that are conducted to identify the needs of different stakeholders. Therefore requirements analysis means to analyse, document, validate and manage software or system requirements. High-quality requirements are documented, actionable, measurable, testable, traceable, helps to identify business opportunities, and are defined to a facilitate system design. After the extensive analysis

of the problems in the system, we are familiar with the requirement that the current system needs. The system requirements are categorized into the functional and non-functional requirements. These requirements are listed below:

A. Hardware requirements

- 1) Minimum 4GB RAM.
- 2) 200 MB of free Hard Disk space.
- 3) Browser: Chrome(v49), Internet Explorer (v10) or higher.
- 4) Processor: 3GHz or higher.

B. Software requirements :-

- 1) Operating System : Windows 7 and above.
- 2) UiPath Studio.
- 3) UiPath Robots.
- 4) UiPath Orchestra.

C. Functional requirement

Functional requirement are the functions or features that must be included in any system to satisfy the business needs and must be acceptable to the users. Based on this, the functional requirements that the system must require are as follows:

- 1) Name and Address of the Recipient: The Name of Recipient is required to identify the person who has initiated the transaction. The name should be spelled correctly in order to avoid confusion. We require physical address of the customer to deliver any replacement or something else.
- 2) Sold By: This feature describes the seller of the product. Basically it answers the query from where is the product shipped and which outlet is responsible for shipping.
- 3) Total Amount: This field gives us the final amount of the product.
- 4) Order Number: This field describes the order number or order ID of the product. This field can immediately help the seller outlet to identify all the product details.
- 5) Order Date: This field gives us the Date when the given product was purchased.
- 6) Invoice Number: This field shows us the Receipt (Invoice) number of the transaction. This number is

useful in servicing centres to identify the product warranty or guarantee period.

7) Invoice Date: Date at which the invoice was printed.

8) Product Description: This field gives us the information about either products or services including prices and quantities. Often includes standard product description and inventory number.

B. Non-functional requirement :-

Non-functional requirement is a description of features, characteristics and attribute of the system as well as any constraints that may limit the boundaries of the proposed system. The non-functional requirements are essentially based on the performance, information, economy, control and security efficiency and services. Based on these the non-functional requirements are as follows:

1) Security:- Security requirements ensure that the software is protected from unauthorized access to the system and its stored data. It considers different levels of authorization and authentication across different user's roles. For instance, data privacy is a security characteristic that describes who can create, see, copy, change, or delete information. Security also includes protection against viruses and malware attacks.

2) Reliability:- Reliability defines how likely it is for the software to work without failure for a given period of time. Reliability decreases because of bugs in the code, hardware failures, or problems with other system components. To measure software reliability, you can count the percentage of operations that are completed correctly or track the average period of time the system runs before failing.

3) Performance:- Performance is a quality attribute that describes the responsiveness of the system to various user interactions with it. Poor performance leads to negative user experience. It also jeopardizes system safety when it's is overloaded.

4) Availability:- Availability is gauged by the period of time that the system's functionality and services are available for use with all operations. So, scheduled maintenance periods directly influence this parameter. And it's important to define how the impact of maintenance can be minimized. When writing the availability requirements, the team has to define the most

critical components of the system that must be available at all time.

5) Scalability:- Scalability requirements describe how the system must grow without negative influence on its performance. This means serving more users, processing more data, and doing more transactions. Scalability has both hardware and software implications. For instance, you can increase scalability by adding memory, servers, or disk space. On the other hand, you can compress data, use optimizing algorithms, etc.

IV. DESIGN PROCESS

Automated invoice processing helps to save a serious amount of time and money, compared with manual paper invoice processing, creating efficiencies and increasing the accuracy of captured data.

A. PROCESS DIAGRAM

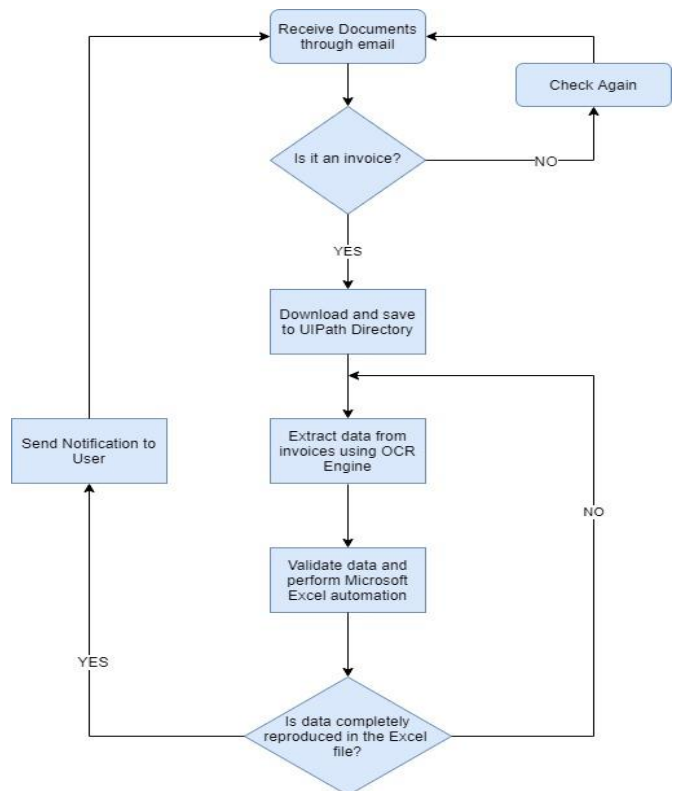


Fig 2. Process Diagram of Invoice Processing Flow

This service enables your company to:

- 1) Capture your invoices from their original source e.g. fax, email, electronic image or paper copy

- 2) Automatically extract key information from your invoices, using clever software for import into your financial systems
- 3) Validate your invoices and checking your invoice data against data files e.g. supplier name, amount and Purchase Order number
- 4) Match and check the purchase order system to make sure the invoice matches the original order

B. FLOW DIAGRAM

Once the data is extracted or captured from the invoice the data is sent into the system for automatic matching against the purchase order. This process can compare just the invoice data with that shown on the purchase order or be expanded to include a deeper level that looks at the receiving documents. Workflow steps can be configured such that the responsible person will then receive an email alert so that he or she can approve the invoice. If there are other people involved in the approval workflow, email alerts to them will also be automatically generated. The typical workflow is a four-step process beginning with:

1. Import of the images through scanning or email,
 2. Identification of the vendor and business unit associated with the invoice,
 3. Data extraction, and
 4. Export of the extracted data and images then receive an email alert so that he or she can approve the invoice.
- If there are other people involved in the approval workflow, email alerts to them will also be automatically generated. The typical workflow is a four-step process beginning with:

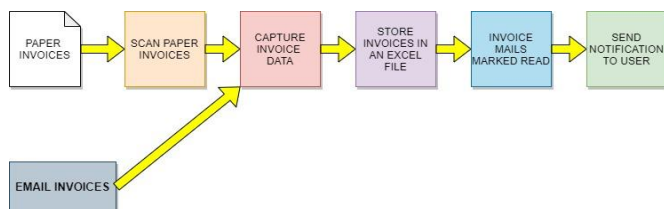


Fig 3. Illustration of Invoice Processing Flow.

V. DESIGN AND IMPLEMENTATION

Regardless the size of your organization, financial processes play an important role in the overall functionality. These days, most people are skeptical regarding the capabilities and success rates of RPA and invoice processing Companies, in a nutshell, receive an invoice from their suppliers sometimes, and usually arrive in a plethora of different formats, including a PDF email attachment, as a paper copy, or fax. In addition, the finance department of an organization is majorly responsible for manually transferring the data from a multitude of invoice formats into the database of the company and dealing with any incongruity.

Without a shadow of a doubt, these things really make it difficult and allow a business to carry out the overall process manually and on a faster rate.

Since the data in these invoices are not structured in a proper format, this manual transfer process is time-consuming and there are chances that a mistake could be made.

In our proposed system DataBot will perform Invoice automation that would free up back office finance/procurement teams to focus on higher value added tasks.

Invoice automation allows straight through processing (no human interaction) most of the time for the entire invoice process. Invoice automation involves:

A. Monitoring For Invoices

Invoices arrive in companies as PDFs, image files and increasingly rarely as hard copy documents.

- 1) For digital invoices, an RPA bot or a simple email automation tool can flag emails with invoices and forward them for data extraction. Some companies use a dedicated email address for invoices to further simplify invoice monitoring.
- 2) For hard-copy invoices, companies are switching to using a single address to centralize invoice scanning.

B. Invoice Capturing :-

- 1) Extracting relevant details (e.g. bank account, ordered item) from the invoice. If software does not have confidence in the results, it is sent to employees for a manual check
- 2) While digitization helped automate numerous processes, mostly rule based software was used in digitization. Invoice capture software is different. Invoice capture involves both reading the invoice text with Optical Character Recognition (OCR) and understanding its context with machine learning.

C. Evaluating Invoice Against Order Records And Other Criteria :-

Evaluating invoice against order records and other criteria to ensure that the payment is indeed a valid one. Evaluations include

- 1) cross-checking invoice against purchase orders
- 2) cross-checking invoice for duplicity
- 3) using working capital optimization policies to decide payment time
- 4) using limits to decide whether to manually process invoice. Invoices that are abnormally large compared to a suppliers' usual invoices may need to be manually verified to ensure that wrong payments are not done

D. Recording Invoice Related Information in System :-

- 1) Using intelligent OCR and natural language processing capabilities, software robots are able to read out the information that is visible on the invoice.
- 2) After robots extract the key information from each invoice, they use their credentials to open the company's database or enterprise resource planning system, if not already open. The robots then start processing the invoices one-by-one by transferring over the relevant invoice information.

E. Email Notification

- 1) After successfully registering each invoice, the software robots are then able to send posting notifications in the form of emails to the responsible employee or to the vendor in question.

- 2) An email is also sent to the responsible party in case of an exception

VI. TECHNOLOGIES USED

A. UIPATH STUDIO :-

UiPath Studio is a complete solution for application integration, and automating third-party applications, administrative IT tasks and business IT processes. One of the most important notions in Studio is the automation application.

An application is a graphical representation of a business process. It enables you to automate rule-based processes, by giving you full control of the execution order and the relationship between a custom set of steps, also known as activities in UiPath Studio. Each activity consists of a small action, such as clicking a button, reading a file or writing to a log panel.

The main types of supported workflows are:

- 1) Sequences - suitable to linear processes, enabling you to smoothly go from one activity to another, without cluttering your workflow.
- 2) Flowcharts - suitable to a more complex business logic, enabling you to integrate decisions and connect activities in a more diverse manner, through multiple branching logic operators.
- 3) State Machines - suitable for very large workflows; they use a finite number of states in their execution which are triggered by a condition (transition) or activity.
- 4) Global Exception Handler - suitable for determining the workflow behavior when encountering an execution error, and for debugging processes.

B. E-MAIL ACTIVITIES PACKAGE :-

- 1) The Mail Activities Pack is designed to facilitate the automation of any mail-related tasks, covering various protocols, such as IMAP, POP3 or SMTP. UiPath also features activities that are specialized for working with Outlook and Exchange.

VII. RESULTS

2) Activities such as Save Mail Message and Save Attachments are not intended to be used with certain mail protocols. Instead, they save the MailMessage object variable retrieved from activities such as Get POP3 Mail Message to a specified folder on the current machine

C. PDF ACTIVITIES PACKAGE

1) The PDF pack contains activities designed to extract data from PDF and XPS files and store it into string variables. The data can be extracted from the entire document or from a range of pages specified under the Range property found in each of the activities.

2) In the case of scanned documents, data extraction can also be achieved by using OCR-based activities, Read PDF with OCR and Read XPS With OCR. To select one of the three OCR engines specific to UiPath, Google OCR, Microsoft OCR and Abbyy OCR, simply drop the engine in the body of the activity.

D. EXCEL ACTIVITIES PACKAGE :-

1) The Excel activities package aids users to automate all aspects of Microsoft Excel, as we know it is an application intensely used by many in all types of businesses.

2) It contains activities that enable you to read information from a cell, columns, rows or ranges, write to other spreadsheets or workbooks, execute macros, and even extract formulas. You can also sort data, color code it or append additional information.

E. UIPATH ROBOT :-

1) The Robot is an execution agent, meaning that you have to provide it with the automation applications you want it to run.

2) After creating an automation application in Studio, it needs to be published locally or to Orchestrator. Once an application is published, you can send it to the Robot machine and start executing it.

3) This is populated by default as follows:

- i) When NOT connected to Orchestrator
- ii) When connected to Orchestrator - the default Orchestrator feed

Following are the results we obtained after performing a series of test with real time invoices. The accuracy we obtained is 100%.

1. This is the Sample copy of an invoice which we have used for testing.

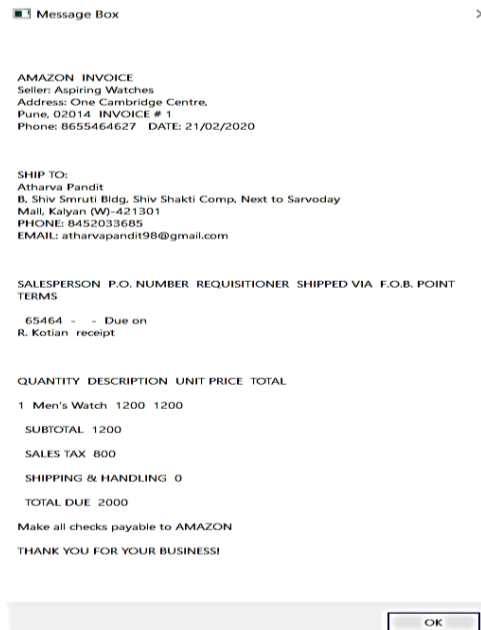


Fig 4: Sample copy of an invoice

2. All Invoices sent to RPA databot through email is successfully downloaded to a dedicated folder which is assigned during development. The databot never re-download the invoice which has been already downloaded previously.

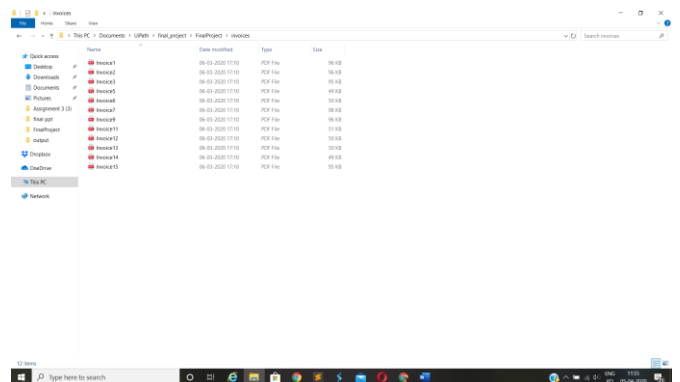


Fig 5: Successful download and storing of invoices in the dedicated folder

3. After successfully downloading the invoices the datobot correctly fetches invoices from that folder one by one. After this it reads all the necessary content from the invoice using an OCR engine then it stores the details in the Excel file with 100% accuracy in their correct places.

	A	B	C	D
	Seller	Seller Address	Client	Client Address
1	Aspiring Watches	One Cambridge Centre,	Atharva Pandit	B, Shiv Smruti Bldg, Shiv Shakti Comp, Nex
2	Laxman Watches	6, LBS Marg	Rasika Patil	5, Shiv Amrut, Gulburga Marg
3	Sanjana Electronics	15, Shil Marg	Sachin Patil	5, Shiv Amrut, Gulburga Marg
4	Chaitanya Steels	1, Nehru Center	Ishant Porel	18, Shakti Mahal
5	Bhaurao Mechanic	15, Nehru Nagar	David Cardozo	1, Amber Apartments
6	Sachin Electricals and Hardware	18, Sardar Patel Marg	Akash Singh	8, Anandiben Society
7	Rathee Electronics	5, Lokmanya Tilak Marg	Raj Natekar	15, Kalpatru Apartments
8	Shivam Eyeglasses	6, Boring Road, Jagannath Plaza	Atharva Pandit	B, Shiv Smruti Bldg, Shiv Shakti Comp, Nex
9	Chaitanya Dresses	7, Akbar Road	Sania Salwekar	F1, Jagannath Apartments,
10	Suresh Handiworks	5, Shivaji Chowk,	Ashwin Thakur	6, Camilia, Vasant Valley,
11	Suresh Handiworks	5, Shivaji Chowk,	Animesh Ghosh	8, Daffodils, Vasant Valley,
12	Sai Jewellers	5, Shree Ram Chowk,	Ramesh Potdar	8, Sunflower, Vasant Valley,
13	Shree Ram Jewellers	5, Shree Ram Chowk,	Amit Salwekar	15, Blossom, Vasant Valley,
14	Krishna Dresses	5, Lal Bahadur Shastri Marg	Jatin Sapru	1, Pragati, Gulburga Marg

Fig 6: Successful data entry of content in the Excel file.

4. After successfully registering each invoice, the software robot is then able to send post notifications in the form of email to the concerned employee or to the vendor in question.

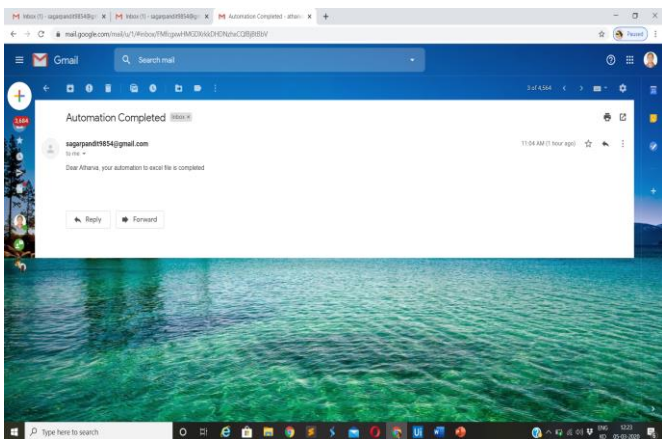


Fig 7: Successfully notifying the concerned employee or vendor.

VIII. CONCLUSION

Automated invoice processing can achieve powerful results for accounts payable departments. Thanks to

technological advancements in robotic process automation and computer vision technologies, invoice processing can eliminate bottlenecks within the AP process and turn the department into the profit center it can be. Automated invoice processing enables touchless automation across the entire accounts payable process, and can transform the business in just months, creating a powerful return on investment. Any organization that receives a large number of vendor invoices on paper can benefit from invoice processing technology. The more data from each invoice that you are hand-keying into your accounting software that more benefits you can get from each page you automate.

IX. REFERENCES

- [1]. Ying Li, Muthu Muthiah, Arindam Routh and Chitra Dorai, "Cognitive Computing in Action to Enhance Invoice Processing with Customized Language Translation", 2017, IEEE International Conference on Cognitive Computing (ICCC), IBM.
- [2]. Hatem Hamza, Yolande Belaid and Abdel Belaid, "A Case-Based Reasoning Approach on Unknown Class Invoice Processing", 2007, IEEE International Conference on Image Processing, France.
- [3]. T. A. Bayer, H. U. Mogg-Schneider and Daimler-Benz, "A Generic System For Processing Invoices", 1997, Fourth International Conference on Document Analysis and Recognition, Germany.
- [4]. Enes Aslan, Ethem Unver, Tugrul Karakaya and Yusuf Sinan Akgul, "An optimization approach for invoice image analysis", 2015, Signal Processing and Communications Application Conference, Turkey

Cite this article as :

Sagar Sahu, Sania Salwekar, Atharva Pandit, Manoj Patil, "Invoice Processing Using Robotic Process Automation", International Journal of Scientific Research in Computer Science, Engineering and Information Technology (IJSRCSEIT), ISSN : 2456-3307, Volume 6 Issue 2, pp. 216-223, March-April 2020. Available at doi : <https://doi.org/10.32628/CSEIT2062106>
Journal URL : <http://ijsrcseit.com/CSEIT2062106>