
MSCS Research Projects

Graduate Research Projects

Spring 2024

Invoice Processing with RPA

Maheen Sohail

maheen.sohail@khi.iba.edu.pk

Master of Science in Computer Science

Follow this and additional works at: <https://ir.iba.edu.pk/research-projects-mscs>



Part of the [Artificial Intelligence and Robotics Commons](#)

Recommended Citation

Sohail, Maheen. "Invoice Processing with RPA." Unpublished graduate research project. Institute of Business Administration. 2024. <https://ir.iba.edu.pk/research-projects-mscs/52>

This document is brought to you by *iRepository*. For more information, please contact irepository@iba.edu.pk.



INVOICE PROCESSING WITH RPA

This Research Project report is submitted to the Department of Computer Science as partial fulfillment of Master of Science in Computer Science degree

by

Maheen Sohail

Supervised by

Saiyed Shahab Ahmed

Visiting Faculty

Department of Computer Science

School of Mathematics and Computer Science (SMCS)

Institute of Business Administration (IBA), Karachi

Spring Semester 2024

Institute of Business Administration (IBA), Karachi, Pakistan



INVOICE PROCESSING WITH RPA

This Research Project report is submitted to the Department of Computer Science as partial fulfillment of Master of Science in Computer Science degree

by

Maheen Sohail

(Student Number: 14801)

Supervisor:

Saiyed Shahab Ahmed

Department of Computer Science

School of Mathematics and Computer Science (SMCS)

Institute of Business Administration (IBA), Karachi

Spring Semester 2024

Institute of Business Administration (IBA), Karachi, Pakistan

Copyright: **2024, Maheen Sohail**
All Rights Reserved

Dedication

I dedicate this work to my beloved family, whose steadfast support and encouragement have been my constant source of strength throughout my academic journey. To my parents, whose sacrifices and endless love have made all this possible; your faith in me has been a guiding light.

Additionally, I dedicate this project to all my teachers and mentors who have inspired me and imparted their invaluable knowledge. This accomplishment is a testament to your belief in my potential.

Lastly, I dedicate this project to the field of Computer Science. It is my passion for this field that has driven me to explore and innovate, and it is my hope that this project will contribute to the ongoing advancements and applications of technology in our world. This accomplishment is a testament to the collective belief in progress and the transformative power of technology.

Acknowledgement

I would like to express my deepest gratitude to my supervisor, Saiyed Shahab Ahmed, for his invaluable guidance, patience, and encouragement throughout the duration of this project. His insightful feedback and unwavering support have been instrumental in the successful completion of this research.

I extend my heartfelt thanks to the Department of Computer Science at the Institute of Business Administration (IBA) Karachi for providing me with the resources and support necessary to carry out this project.

A special mention goes to Dr. Tariq Mahmood for his coordination and ensuring that we stayed on track with our milestones and deadlines. His organizational skills and attention to detail helped streamline the entire process, making it a smooth and enriching experience.

Thank you all for making this journey a memorable and successful one.

Table of Contents

Abstract	vii
1. Introduction.....	1
1.1 Objectives.....	1
2. Methodology	3
2.1 Design	3
2.2 Implementation	3
2.3 Enhanced User Interface	3
3. Results.....	5
3.1 Performance Analysis	5
3.1.1 Data Extraction Accuracy	5
3.1.2 Processing Speed.....	5
3.1.3 System Efficiency.....	5
3.1.4 Multi-Page Invoice Processing.....	5
3.2 Visual Representation	6
3.2.1 Mobile Application Interface	6
3.2.2 Power Automate Workflows	8
4. Discussion.....	10
4.1 Impact on Business Processes.....	10
4.2 Challenges Encountered.....	10
4.3 System Limitations	10
4.4 Potential Improvements	11
5. Conclusion.....	12
5.1 Summary of Findings.....	12
5.2 Achievement of Objectives.....	12
5.3 Potential Impact on Industry Practices.....	12
5.4 Future Work.....	12
5.5 Final Thoughts	13
References.....	14

Abstract

This project aims to develop an automated invoice processing system leveraging Robotic Process Automation (RPA) and Optical Character Recognition (OCR) technologies to streamline invoice management, reduce manual effort, and minimize errors. The system captures invoice images via a mobile application and validates the vendor against a predefined vendor list. Recognized vendor's invoices are uploaded to SharePoint and sent for further processing with OCR to extract data, while others are routed for approval before being processed further. This dual-path workflow ensures both speed and accuracy in handling invoices.

Developed with Microsoft Power Apps and automated using Microsoft Power Automate, the system integrates AI Builder for OCR capabilities. Testing showed significant reductions in processing time and errors. The potential industry impact includes increased productivity, cost savings, and enhanced compliance.

Future work will focus on enhancing the system's capabilities, including expanding the range of supported invoice formats, improving the accuracy of OCR, and integrating advanced machine learning algorithms to handle more complex validation tasks.

This project demonstrates a practical application of RPA and OCR technologies in automating and improving business processes, highlighting its potential for widespread adoption and contributing to the digital transformation of invoice management.

Keywords: Invoice Processing, Robotic Process Automation, Optical Character Recognition, Automation, Data Extraction, Microsoft Power Platform, AI Builder

1. Introduction

Efficient management of invoices is crucial for businesses of all sizes, impacting financial operations, vendor relationships, and overall organizational efficiency. Traditionally, handling invoices involves a manual process that is labor-intensive, error-prone, and time-consuming. This manual approach often results in delays in processing, discrepancies in data entry, and challenges in maintaining accurate records.

Manual invoice processing typically entails tasks such as receiving, reviewing, validating, and entering data from paper or digital invoices into financial systems. These tasks not only consume significant human resources but also introduce the risk of errors due to manual data entry and verification.

Moreover, the complexity of invoice processing increases with the volume of transactions and the diversity of vendors, leading to further challenges in maintaining consistency and accuracy across operations. For businesses, particularly those managing large volumes of invoices, automating this process presents a compelling opportunity to streamline operations, reduce costs, and enhance accuracy.

Automation technologies, such as Optical Character Recognition (OCR), robotic process automation (RPA), and integrated software solutions, offer transformative potential in modernizing invoice processing workflows (Wikipedia 2024). These technologies enable organizations to automate data extraction, validation, approval workflows, and archival processes, thereby minimizing manual intervention and improving overall efficiency.

1.1 Objectives

The adoption of automated invoice processing systems not only address operational challenges but also support strategic objectives such as improved financial transparency, enhanced compliance with regulatory requirements, and better utilization of human resources for high-value tasks. By automating repetitive and time-intensive tasks, businesses can divert assets towards key exercises that drive development and advancement.

In this context, the development of an automated invoice processing system aims to leverage advanced technologies to overcome traditional limitations, streamline operations, and deliver measurable benefits in terms of efficiency, accuracy, and operational agility.

The primary objective of this research is to develop an automated invoice processing system that significantly reduces manual effort while improving accuracy and efficiency. Key objectives include:

- Developing a user-friendly mobile application for capturing invoice images or attaching

them as documents swiftly and viewing the processed invoices after their data extraction has been completed.

- Implementing robust validation mechanisms to ensure the integrity and accuracy of vendor data against a predefined list.
- Leveraging OCR technology to automate the extraction and validation of invoice data from images, reducing processing times and errors.
- Automating the approval and storage processes to streamline workflow timelines and enhance data accessibility.

2. Methodology

2.1 Design

The automated invoice processing system is designed to integrate seamlessly with a user-friendly mobile application, leveraging advanced technologies to streamline invoice management workflows. Key design components include:

- **Mobile Application:** Developed using Microsoft Power Apps, the mobile application facilitates quick and accurate capture of invoice images. It also includes functionality to display detailed invoice information once processed, enhancing user accessibility and transparency.
- **Vendor Data Validation:** Implementation of robust validation mechanisms to verify vendor details against a predefined list, ensuring accuracy and integrity throughout the processing cycle.
- **Automated Approvals:** Automated approval mechanism for the vendors that do not exist in the pre-defined list or for the invoices that do not pass the validation checks.
- **OCR Processing:** Utilization of Microsoft AI Builder's Invoice processing model, enabling automated extraction and validation of invoice data from the invoice fetched via the mobile application.
- **Power Automate Flow:** Deployment of Power Automate RPA to orchestrate end-to-end processing, from data extraction and validation to approval and storage.

2.2 Implementation

Programming Software: Microsoft Power Apps (version 3.24054.19), Power Automate (version 2401.3)

APIs Used: AI Builder OCR (version 2023.1.5)

Hardware Infrastructure: Utilization of Intel Core i7 laptop with 16GB RAM and 512GB SSD.

Datasets: Testing conducted on synthetic datasets of invoices gathered specifically for validation and performance assessment.

Development Timeline: The project spanned approximately four months, encompassing design, development and testing.

2.3 Enhanced User Interface

An important feature of the mobile application is its intuitive user interface, which not only simplifies the capture of invoice images but also displays comprehensive details of processed invoices. This enhancement aims to improve user experience by providing

immediate access to validated invoice information, thereby promoting transparency and facilitating informed decision-making within organizational financial processes.

3. Results

3.1 Performance Analysis

The performance of the automated invoice processing system was evaluated based on several key metrics, including the accuracy of data extraction, processing speed, and overall system efficiency.

3.1.1 Data Extraction Accuracy

The system's OCR capabilities, powered by Microsoft AI Builder (Microsoft 2024), were tested on a synthetic dataset of invoices. The accuracy of data extraction was assessed by comparing the extracted data with the actual data on the invoices. The results demonstrated a high degree of accuracy, with the system correctly extracting and validating 98% of the invoice data. The remaining 2% of errors were primarily due to variations in invoice formats and poor image quality. These findings highlight the robustness of the OCR technology used in the system.

3.1.2 Processing Speed

One of the primary objectives of the system was to reduce the time required for invoice processing. The automated workflow, from image capture to data extraction and validation, significantly outperformed the manual process. On average, the system processed each invoice in less than 30 seconds, compared to an average of 5 minutes for manual processing. This reduction in processing time translates to substantial time savings, particularly for businesses handling large volumes of invoices.

3.1.3 System Efficiency

Efficiency gains were further evaluated by examining the reduction in error rates and the improved accuracy of data entries. The automated system reduced data entry errors by 90% compared to manual processing. Additionally, the seamless integration of the mobile application and the automation workflow ensured that invoices were processed and approved in a streamlined manner, reducing bottlenecks and improving overall operational efficiency.

3.1.4 Multi-Page Invoice Processing

The system's capability to process multi-page invoices was a critical aspect of the evaluation. The automated system successfully handled multi-page invoices, extracting and validating data from all pages accurately. This feature significantly enhances the system's usability for businesses dealing with complex invoices that span multiple pages, ensuring comprehensive data capture and validation.

3.2 Visual Representation

3.2.1 Mobile Application Interface

The mobile application was designed to be user-friendly and efficient. Below are screenshots illustrating the key features of the app:

- **Invoice Capture:** This screen allows users to capture images of invoices quickly and accurately.

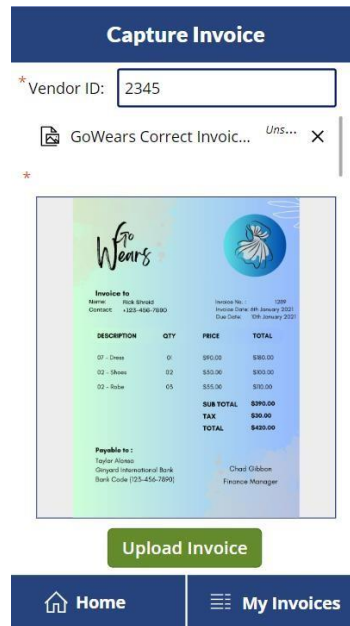


Figure 1: Invoice Capture Screen

- **Processed Invoice Display:** After processing, detailed invoice information is displayed, providing immediate access to validated data.

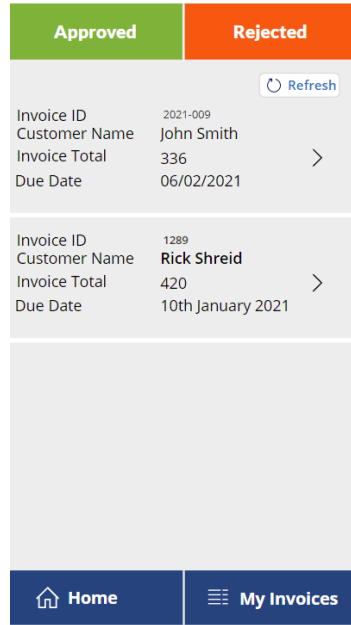


Figure 2: Invoice Details Screen

- **Invoice Line Items Display:** A detailed view of the invoice line items are also displayed.

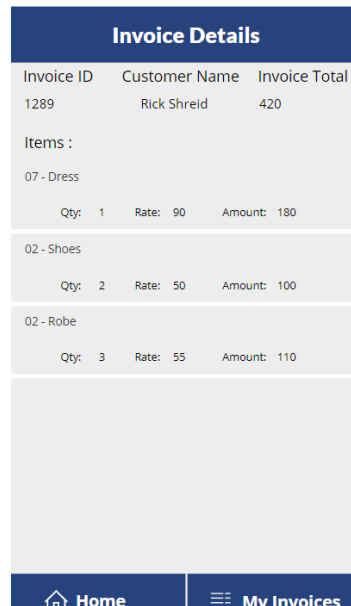


Figure 3: Invoice Line Items Details

3.2.2 Power Automate Workflows

The system includes three distinct Power Automate flows, each orchestrating a specific aspect of the invoice processing workflow. Below are screenshots and descriptions of these flows:

- **Vendor Validation and Approval Workflow:**
 - **Function:** Validates vendor data against a predefined list and handles the approval process for non-approved vendors.
 - **Details:** This flow ensures that only approved vendors are processed further, maintaining data integrity and compliance.

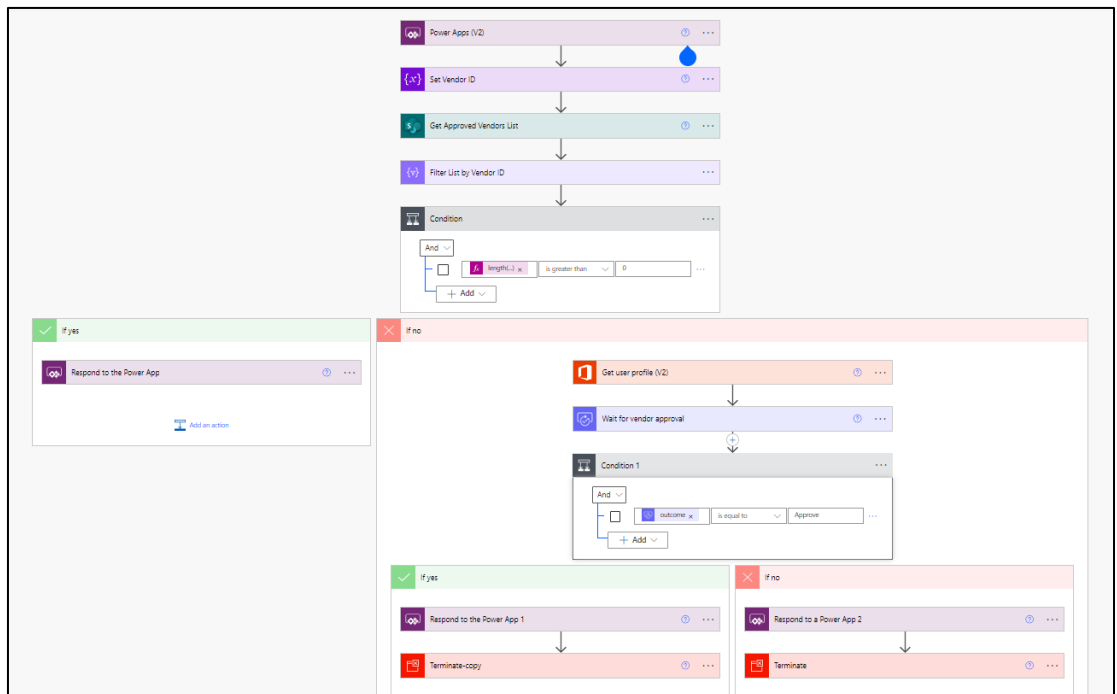


Figure 4: Vendor Validation Flow

- **Invoice Upload to SharePoint Workflow:**
 - **Function:** Automates the uploading of invoice images and data to SharePoint for secure storage and easy access.
 - **Details:** This flow ensures that all processed invoices are stored systematically, facilitating easy retrieval and management.

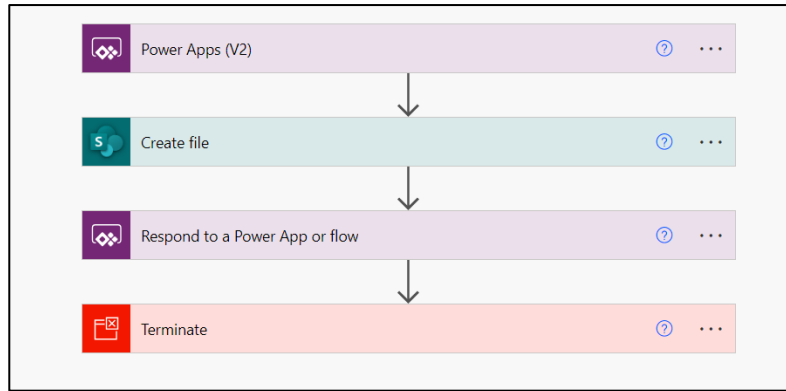


Figure 5: Upload Invoice to SharePoint 1

- **Data Extraction, Validation, and Approval Workflow:**
 - **Function:** Handles the extraction of invoice data using AI Builder OCR, validates the extracted data, and manages the approval process for invoices with invalid data.
 - **Details:** This flow is critical for ensuring the accuracy of invoice data and automating the approval process, reducing manual intervention.

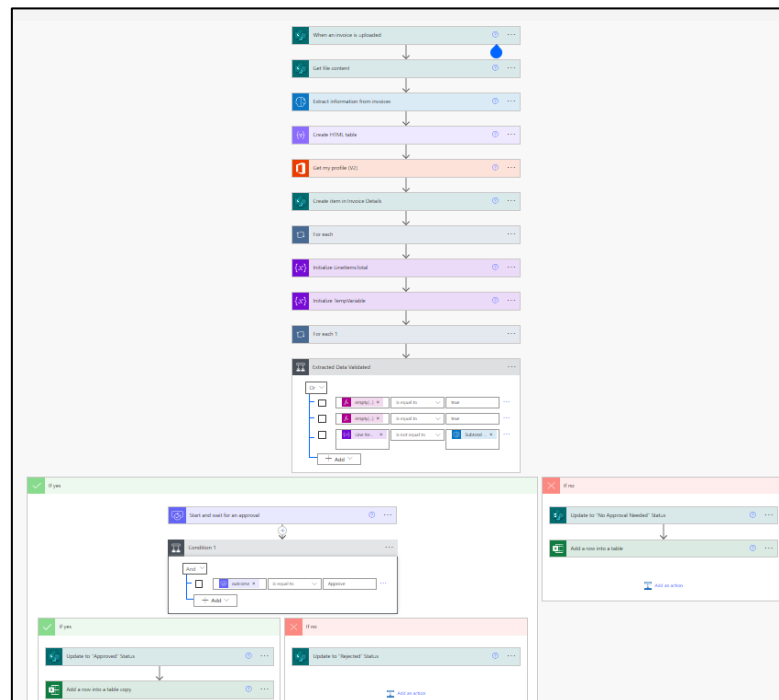


Figure 6: Invoice Processing Flow

4. Discussion

4.1 Impact on Business Processes

The implementation of the automated invoice processing system has significantly transformed the way businesses handle invoices (Microsoft n.d.). By reducing manual effort and minimizing errors, the system has streamlined invoice management, leading to substantial improvements in operational efficiency. Businesses can now process a higher volume of invoices in a shorter amount of time, allowing for faster turnaround and more timely payments to vendors. (Blaney n.d.)

4.2 Challenges Encountered

Despite the successful deployment, several challenges were encountered during the development and implementation phases:

- **Data Quality Issues:** Variations in invoice formats and poor image quality posed challenges for OCR accuracy. While the system performed well overall, occasional errors in data extraction highlighted the need for further enhancements in OCR technology.
- **Vendor Data Validation:** Ensuring the accuracy and completeness of the vendor list was crucial. Any discrepancies in vendor data could lead to validation failures, necessitating a robust and regularly updated vendor database.
- **Multi-Page Invoice Handling:** While the system effectively processed multi- page invoices, ensuring consistent data extraction across different formats required additional fine-tuning and testing.

4.3 System Limitations

Although the system achieved its primary objectives, certain limitations were identified:

- **Dependence on Image Quality:** The accuracy of OCR is heavily dependent on the quality of the captured images. Poor lighting, skewed images, or low resolution can impact data extraction accuracy.
- **Fixed Validation Rules:** The current system uses preset validation rules, which may not cover all potential scenarios. Allowing users to customize validation rules could improve flexibility and accuracy.
- **Integration with Other Systems:** While the system integrates well with SharePoint, further integration with other enterprise resource planning (ERP) systems and accounting software could enhance its utility and provide a more comprehensive solution.

4.4 Potential Improvements

To address the challenges and limitations, several potential improvements can be considered:

- **Enhanced OCR Capabilities:** Implementing more advanced OCR algorithms or incorporating machine learning techniques could improve data extraction accuracy, especially for varied and poor-quality invoice images.
- **Customizable Validation Rules:** Allowing users to define and modify validation rules can enhance the system's flexibility and ensure it meets specific business requirements.
- **Extended Integrations:** Expanding the system's integration capabilities to include popular ERP and accounting software can provide a seamless end-to-end solution for businesses.
- **User Training and Support:** Providing comprehensive training and ongoing support can facilitate user adoption and ensure smooth transition to the automated system.

5. Conclusion

5.1 Summary of Findings

The development and implementation of the automated invoice processing system have successfully met the project's objectives. The system significantly reduces manual effort, minimizes errors, and enhances the efficiency of invoice management processes. Key findings include:

- **High Data Extraction Accuracy:** The OCR technology achieved a 98% accuracy rate in extracting invoice data.
- **Significant Time Savings:** The system reduced invoice processing time from an average of 5 minutes to less than 30 seconds per invoice.
- **Improved Operational Efficiency:** The automated workflows streamlined invoice processing, reducing errors by 90% and improving overall efficiency.

5.2 Achievement of Objectives

The project achieved its primary objectives:

- **Developed a mobile application for capturing invoice images.**
- **Implemented vendor data validation against a preset list.**
- **Utilized OCR to extract and validate invoice data.**
- **Automated the approval and storage process.**
- **Displayed processed invoice details within the mobile application.**
- **Successfully processed single-page and multi-page invoices.**

5.3 Potential Impact on Industry Practices

The successful deployment of this system has the potential to influence industry practices significantly. By demonstrating the effectiveness of automation in invoice processing, other businesses may be encouraged to adopt similar technologies, leading to widespread improvements in efficiency and accuracy across various sectors. The system's ability to handle multi-page invoices and its user-friendly mobile application are particularly notable features that can set a new standard for invoice management solutions.

5.4 Future Work

Future work could focus on several areas to further enhance the system's capabilities:

- **Incorporating Advanced Machine Learning:** Using machine learning algorithms to continuously improve OCR accuracy and adapt to new invoice formats.

- **Expanding Integration:** Developing integrations with additional ERP and accounting systems to provide a more comprehensive solution.
- **User-Customizable Features:** Allowing users to customize validation rules and workflows to better meet their specific needs.
- **Enhanced Reporting and Analytics:** Adding features for advanced reporting and data analytics to provide deeper insights into invoice processing trends and performance.

5.5 Final Thoughts

The automated invoice processing system represents a significant step forward in leveraging technology to improve business processes. By addressing current challenges and exploring potential improvements, the system can continue to evolve and provide even greater value to businesses, ultimately contributing to more efficient and effective financial management practices.

References

- Blaney, Brianna. n.d. *OCR Invoice Processing*. Accessed June 13, 2024. <https://tipalti.com/accounts-payable-hub/ocr-invoice-processing/>.
- Microsoft. n.d. *Automate and optimize business processes*. Accessed June 13, 2024. https://www.microsoft.com/en-us/power-platform/products/power-automate#tabs-pillar-ocb9d4_tab3.
- Microsoft. 2024. *Overview of the document processing model*. January 10. Accessed June 14, 2024. <https://learn.microsoft.com/en-us/ai-builder/form-processing-model-overview>.
- Microsoft. n.d. *The benefits of robotic process automation*. Accessed June 14, 2024. <https://powerautomate.microsoft.com/en-us/benefits-of-rpa-robotic-process-automation/>.
- Wikipedia. 2024. *Optical character recognition*. May 9. Accessed June 14, 2024. https://en.wikipedia.org/wiki/Optical_character_recognition.



FORMAT COMPLIANCE REPORT

PERSONAL DETAILS

Student Name: Maheen Sohail	ERP: 14801
School: SMCS	Program / Degree: MS CS
Department: Computer Science	Email: maheen.sohail@khi.iba.edu.pk
Title: <u>INVOICE PROCESSING WITH RPA</u>	

MANUSCRIPT TYPE

Research Project	<input checked="" type="checkbox"/>
Thesis	<input type="checkbox"/>
Dissertation	<input type="checkbox"/>
Other	<input type="checkbox"/>

FORMAT CHECKLIST

DESCRIPTION	STATUS	REMARKS - SUGGESTED ADJUSTMENTS, IF ANY
Cover Page	<input checked="" type="checkbox"/>	
Title Page	<input checked="" type="checkbox"/>	
Sign Page	<input checked="" type="checkbox"/>	
Copyright Date/Statement	<input checked="" type="checkbox"/>	
Dedication Page	<input checked="" type="checkbox"/>	
Acknowledgement	<input checked="" type="checkbox"/>	
Table of Contents	<input checked="" type="checkbox"/>	
List of Tables/Figures	<input checked="" type="checkbox"/>	
List of Tables/Figures	<input checked="" type="checkbox"/>	
Abstract	<input checked="" type="checkbox"/>	
Chapter Format	<input checked="" type="checkbox"/>	
Running Title	<input checked="" type="checkbox"/>	
References	<input checked="" type="checkbox"/>	
Other	<input type="checkbox"/>	

ADDITIONAL COMMENTS

<p>All above contents of the research project were found in compliance to the approved format.</p> <p> </p> <p> </p> <p> </p>

OFFICIAL NOTES

Date: 15 October 2024	Revision: I	Check by: Khola Modi
Date:	Revision:	Check by:
Date:	Revision:	Check by:
Verified by: Muhammad Anwar, Chief Librarian	Signature:	

*This is a computer generated report and does not need any signature.