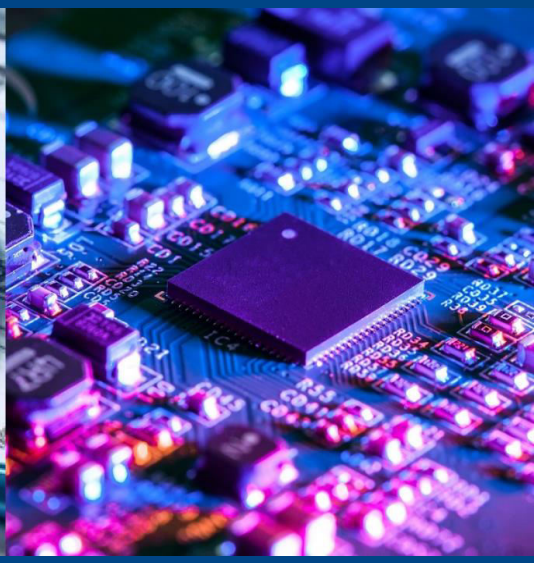
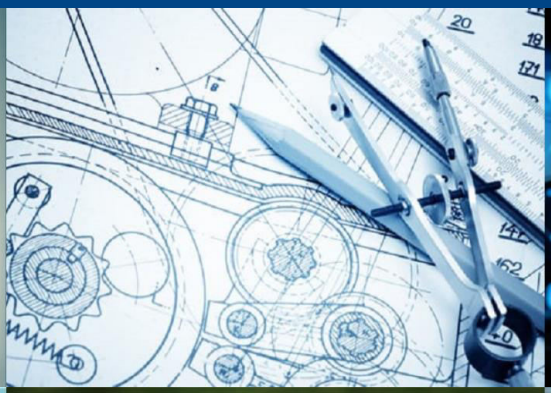
ISSN: 2582-7219

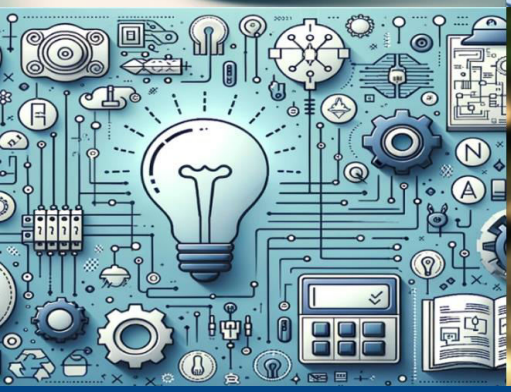




International Journal of Multidisciplinary Research in Science, Engineering and Technology

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)



Impact Factor: 9.864

Volume 9, Issue 6, June 2026

© 2026 IJMRSET | Volume 9, Issue 6, June 2026| DOI:10.15680/IJMRSET.2026.0906028

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 9.864| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Automated Invoice Processing and GST Data Extraction System

R. Nithish Kumar [1], A. Mary Joycy [2]

PG Scholar, Department of Master of Computer Applications, R.V.S College of Engineering, Dindigul, Tamil Nadu, India[1]

Assistant Professor, Department of Master of Computer Applications, R.V.S College of Engineering, Dindigul, Tamil Nadu, India[2]

ABSTRACT: Organizations that handle large volumes of GST invoices often face difficulties in processing and managing invoice data efficiently. Manual extraction of invoice details is time-consuming, prone to human errors, and can lead to inconsistencies in financial records. Additionally, variations in invoice formats and document quality make accurate data extraction a challenging task, resulting in reduced productivity and increased operational effort.

To overcome these challenges, the Automated Invoice Processing and GST Data Extraction System has been developed to automate the extraction, validation, and management of GST invoice information. The system intelligently processes invoice documents in various formats and extracts important details such as invoice number, invoice date, GSTIN, vendor information, GST rate, and taxable amount. The extracted data is presented to users for verification and correction before being stored, ensuring both accuracy and reliability.
The system provides a centralized platform for invoice management, enabling users to upload documents, review extracted information, maintain invoice records, search historical data, and generate structured outputs for further processing. Administrative functionalities are also incorporated to facilitate user management and secure access control.

By automating invoice processing and GST data extraction, the system significantly reduces manual workload, minimizes data-entry errors, improves processing speed, and enhances the overall efficiency of financial document management. The proposed solution offers a reliable and scalable approach for organizations seeking to streamline GST invoice handling and improve operational productivity.

I. INTRODUCTION

In today's digital era, organizations generate and process a large number of invoices as part of their daily business operations. Managing these invoices manually requires significant time and effort, often leading to errors, inconsistencies, and delays in data processing. GST invoices contain critical information such as invoice numbers, dates, GSTIN details, vendor information, tax rates, and taxable amounts, which must be accurately recorded and maintained for financial and regulatory purposes. The Automated Invoice Processing and GST Data Extraction System is developed to address these challenges by automating the extraction and management of GST invoice data. The system intelligently processes invoice documents, extracts relevant information, and presents the results for user verification before storage. By reducing manual intervention and improving data accuracy, the system enhances operational efficiency, streamlines invoice management, and provides a reliable solution for handling GST-related financial documents.

II. EXISTING SYSTEM

In the existing system, GST invoice processing is primarily performed through manual data entry or conventional invoice management tools. Users are required to review invoice documents and manually extract important details such as invoice number, invoice date, GSTIN, vendor information, tax rates, and taxable amounts before storing them in databases or spreadsheets. This process is time-consuming, labor-intensive, and prone to human errors. In organizations handling large volumes of invoices, manual processing can lead to delays, data inconsistencies, and increased operational costs. Additionally, traditional systems often have limited automation capabilities, making it difficult to

© 2026 IJMRSET | Volume 9, Issue 6, June 2026| DOI:10.15680/IJMRSET.2026.0906028

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 9.864| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

efficiently manage, search, and maintain large collections of invoice records. These limitations highlight the need for a more intelligent and automated approach to GST invoice processing and data extraction.

III. PROPOSED SYSTEM

The proposed Automated Invoice Processing and GST Data Extraction System is designed to automate the extraction, validation, and management of GST invoice data. The system intelligently processes invoice documents and extracts important information such as invoice number, invoice date, GSTIN, vendor details, GST rate, and taxable amount. This automation reduces the need for manual data entry and minimizes the possibility of human errors.

The system provides a user-friendly platform where users can upload invoice documents, review the extracted information, and make corrections if required before saving the data. By incorporating a verification mechanism, the system ensures that the stored information is accurate and reliable. All invoice records are maintained in a centralized repository, making data retrieval and management more efficient.

In addition to invoice extraction, the proposed system offers features such as invoice history management, search functionality, secure user authentication, and role-based access control. These features enhance security, improve operational efficiency, and provide a scalable solution for organizations handling large volumes of GST invoices. The system ultimately streamlines invoice processing workflows and supports effective financial document management.

IV. SYSTEM ARCHITECTURE

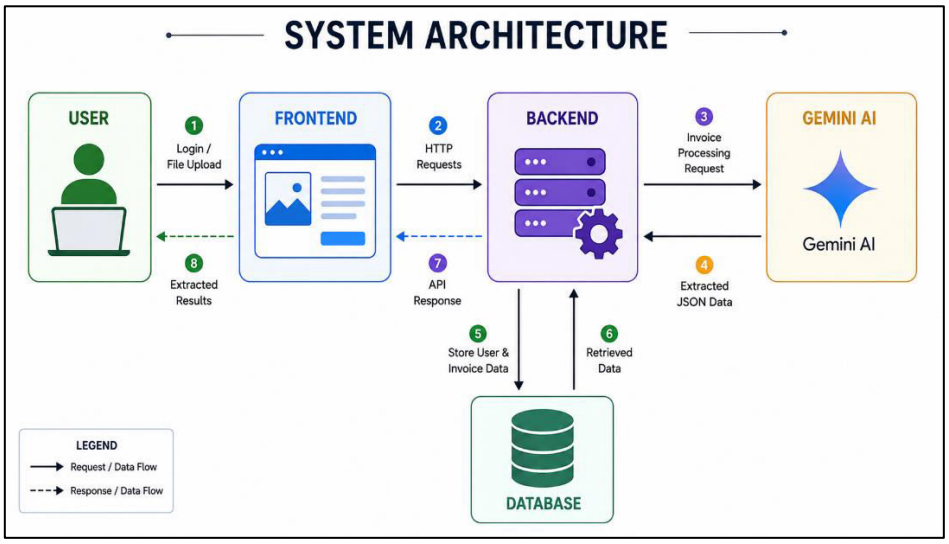


Figure 4.1 ARCHITECTURE DESIGN

The proposed system architecture is designed to automate invoice data extraction through a structured workflow involving the User, Frontend, Backend, Gemini AI, and Database components. The process begins when a user logs into the application and uploads an invoice through the frontend interface. The frontend sends the user request and uploaded file to the backend using secure HTTP communication. The backend validates the request, manages authentication, and forwards the invoice to Gemini AI for intelligent processing. Gemini AI analyzes the document and extracts important invoice details such as invoice number, vendor information, dates, GST details, taxable amounts, and total values, returning the extracted information in a structured JSON format. The backend then stores both the user information and extracted invoice data in the database, ensuring secure and centralized data management. Whenever required, the backend retrieves stored information from the database and sends an API response back to the frontend. Finally, the frontend presents the extracted results to the user in a clear and organized format. This architecture ensures efficient invoice processing, secure data storage, seamless communication between components, and an enhanced user experience through AI-powered document extraction and management

© 2026 IJMRSET | Volume 9, Issue 6, June 2026| DOI:10.15680/IJMRSET.2026.0906028

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 9.864| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

V. MODULE DESCRIPTION

1. Authentication Module:
The Authentication Module manages user registration and login in the Automated Invoice Processing and GST Data Extraction System. It verifies user credentials and ensures that only authorized users can access the system. After successful login, a secure authentication token is generated to validate future requests. The module also implements role-based access control, allowing administrators and regular users to access only the features assigned to their roles. This helps protect sensitive invoice and user data from unauthorized access.

2. Dashboard and Insight Module:
The Dashboard and Insight Module provides users with an overview of important system information. It displays details such as the total number of invoices processed, recent uploads, and extracted records. The dashboard helps users monitor system activities and access important information quickly. It also provides useful insights that assist users in managing and analyzing invoice data efficiently.

3. Invoice Upload and Extraction Module:
The Invoice Upload and Extraction Module allows users to upload GST invoices in formats such as PDF, JPG, JPEG, PNG, and WebP. Once uploaded, the invoice is processed automatically, and the AI service extracts important GST-related information such as invoice number, invoice date, GSTIN, vendor details, GST amount, and taxable value. The extracted data is displayed to the user for verification and is then stored in the database. This module reduces manual effort, improves accuracy, and speeds up invoice processing.

4. History Module:
The History Module stores and displays all previously processed invoices. Users can view invoice details such as vendor name, invoice number, GSTIN, date, and taxable amount. The module also provides search functionality to quickly find specific invoices. Additionally, users can download the extracted JSON data, refresh records, or delete invoices when required. This module helps maintain and manage invoice records efficiently.

5 Admin Module:
The Admin Module is designed for administrators to manage users and monitor system activities. Administrators can view user details, create new accounts, update user information, assign roles, activate or deactivate accounts, and remove users when necessary. The module implements role-based access control to ensure that administrative functions are accessible only to authorized users. This improves system security and helps maintain proper user management.

VI. RESULTS

The Automated Invoice Processing and GST Data Extraction System was successfully developed and tested. The system effectively processes uploaded invoice documents, extracts relevant GST-related information, and presents the extracted data in a structured format. The application provides a user-friendly interface for invoice management, data viewing, and downloading extracted information.

The implemented system successfully performs user authentication, invoice upload, automated data extraction, data storage, and retrieval operations. Extracted invoice details such as invoice number, invoice date, GSTIN, supplier information, customer information, taxable amount, CGST, SGST, IGST, total GST amount, and total invoice value are accurately identified and displayed to the user. The extracted data is automatically converted into a JSON (JavaScript Object Notation) format, enabling structured storage, easy retrieval, and seamless integration with other applications and systems.

The system also provides an option to download the extracted GST data as a JSON file, allowing users to maintain digital records and use the extracted information for reporting, auditing, and further processing. Additionally, the application maintains a history of processed invoices, making invoice tracking and management more efficient.

© 2026 IJMRSET | Volume 9, Issue 6, June 2026| DOI:10.15680/IJMRSET.2026.0906028

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 9.864| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

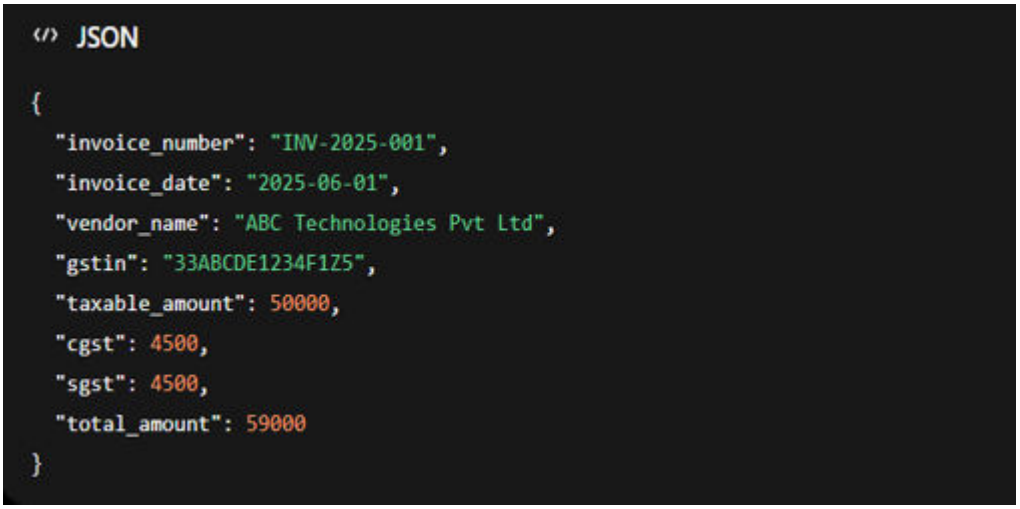


Figure 6.1 JSON GST EXTRACTED DATA

VII. CONCLUSION

This project successfully developed an AI-powered invoice processing and GST data extraction system that automates the extraction and management of invoice data. The system integrates a React frontend, Express.js backend, PostgreSQL database, and Google Gemini's multimodal capabilities to provide an efficient solution without depending on traditional OCR methods.

The application enables users to securely upload invoice documents, automatically extract essential GST-related information, verify and edit extracted data, and maintain a centralized invoice history. Additional functionalities such as role-based authentication, dashboard analytics, search capabilities, and user management improve both security and operational efficiency.

Various testing techniques, including unit testing, integration testing, UI testing, and beta testing, were performed to evaluate the system's functionality and reliability. The results demonstrated that the application performs accurately, securely, and consistently under different usage scenarios.

Overall, the developed system provides a reliable, scalable, and user-friendly approach to intelligent invoice processing. By automating data extraction and reducing manual effort, the system improves productivity, minimizes human errors, and streamlines invoice management operations.

VIII. FUTURE ENHANCE MENT

The developed system successfully automates invoice processing and GST data extraction; however, additional features can be implemented to further improve its functionality, efficiency, and user experience. Enhancements can be introduced to support large-scale business operations, improve data accessibility, and provide more advanced reporting capabilities. These improvements would make the system more flexible, scalable, and suitable for handling increasing volumes of invoice data.

Future enhancements include:

- Bulk Invoice Upload: Enable users to upload and process multiple invoices simultaneously.
- Excel and CSV Export: Allow users to export invoice records in different file formats for reporting and analysis.
- Cloud Storage Integration: Store uploaded files in cloud platforms for better accessibility and backup management.
- Advanced Analytics Dashboard: Provide detailed reports, charts, and invoice trends.

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 9.864| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

- Notification System: Notify users about completed processing tasks and important system activities.
- Performance Optimization: Improve processing speed and overall system efficiency for handling larger datasets.

REFERENCES

[1] Banks, Alex, and Eve Porcello. Learning React: Modern Patterns for Developing React Apps. O'Reilly Media, 2nd Edition, 2020.
[2] Cantelon, Mike, Marc Harter, T.J. Holowaychuk, and Nathan Rajlich. Node.js in Action. Manning Publications, 2nd Edition, 2017.
[3] Hahn, Evan M. Express in Action: Writing, Building, and Testing Node.js Applications. Manning Publications, 2016.
[4] Flanagan, David. JavaScript: The Definitive Guide. O'Reilly Media, 7th Edition, 2020.
[5] Momjian, Bruce. PostgreSQL: Introduction and Concepts. Addison-Wesley Professional, 2nd Edition, 2001.
[6] Merkel, Dirk. "Docker: Lightweight Linux Containers for Consistent Development and Deployment." Linux Journal, Vol. 2014, No. 239, 2014.
[7] Chacon, Scott, and Ben Straub. Pro Git. Apress, 2nd Edition, 2014.
[8] Richardson, Leonard, and Sam Ruby. RESTful Web APIs. O'Reilly Media, 2013.
[9] Sakimura, Nat, John Bradley, and Michael Jones. JSON Web Token (JWT) and OpenID Connect. Internet Engineering Task Force (IETF), RFC 7519, 2015.
[10] Microsoft. Visual Studio Code Documentation. Microsoft Corporation, 2026.

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

INTERNATIONAL JOURNAL OF MULTIDISCIPLINARY RESEARCH

IN SCIENCE, ENGINEERING AND TECHNOLOGY

| Mobile No: +91-6381907438 | Whatsapp: +91-6381907438 | ijmrset@gmail.com |

www.ijmrset.com