Logistics Automation & PDF-to-CSV Processing Platform

1. User Panel

Login

Users can access the platform through a secure login system using email and password authentication. After successful login, the user is redirected to the main dashboard.

2. Dashboard

The dashboard serves as the central interface for users to monitor and manage all activities within the system.

It provides:

- Overview of recent operations
- Summary of shipping requests
- Status of invoices and bills
- Quick navigation to all modules

The dashboard is designed to provide real-time visibility into ongoing and completed processes, allowing users to efficiently manage their workflow.

3. Customer Module

Add Customer

Users can create new customer records by entering required details such as name, contact information, and address.

Manage Customers

This section allows users to:

- View all customers in a structured list
- Edit existing customer details
- Delete customer records
- Search and filter customers

4. Shipping Request Module

Create Shipping Request

Users can create shipping requests by entering:

- Pickup location
- Drop location
- Shipment details

This process ensures that all logistics data is captured in a structured format.

View Past Requests

Users can access historical shipping data, including:

- Previously created requests
- Status tracking
- Shipment history

5. Bill of Lading Module

Create Bill

Users can generate a Bill of Lading (BOL) by associating shipment data with billing information.

Upload Template

The system supports uploading predefined templates to standardize bill generation across different formats.

View Bill of Lading

Users can:

- View all generated bills
- Download documents
- Maintain a record of all transactions

6. Invoice Module

The invoice module allows users to:

- Create invoices linked to shipments
- Edit invoice details
- Delete records if required
- Export invoice data for reporting purposes

7. Location Module

Manage Pickup and Drop Locations

Users can:

- Add new locations
- Update existing entries
- Maintain a centralized database of all operational locations

8. Automation System

The platform includes an automation layer designed to reduce manual effort and increase efficiency.

End-to-End Automation

The system automates:

- Data entry processes
- Workflow execution
- Repetitive operational tasks

Testing Automation

Automated testing ensures system reliability through:

- UI testing
- Regression testing
- Continuous validation of workflows

9. PDF to CSV Conversion Module

This module is responsible for converting raw PDF documents into structured CSV data.

Input

Users upload raw PDF files containing unstructured or semi-structured data.

Processing

The system extracts data using a combination of:

- Text extraction libraries
- OCR techniques for scanned documents
- Pattern recognition using regular expressions

Output

The extracted data is:

- Structured into tabular format
- Converted into CSV files
- Validated for accuracy

---

10. Data Aggregation and Processing

The system processes extracted data to generate meaningful outputs.

Data Processing

- Cleaning and transformation of raw data
- Structuring datasets for analysis
- Aggregation across multiple sources

Data Validation

- Ensures data completeness
- Identifies inconsistencies
- Logs errors for debugging

11. Data Analysis and Reporting

The platform provides analytical capabilities to derive insights from processed data.

Analysis

- Trend identification
- Data filtering and grouping
- Statistical operations

Reporting

- Generation of structured reports
- Exportable datasets
- Visual summaries for decision-making

12. Desktop Application (Tkinter)

A desktop-based interface is provided to interact with the system.

Features

- Form-based data input
- File upload for PDF processing
- Execution of automation workflows
- Integration with backend APIs

The interface is designed to be simple, functional, and efficient for operational users.

13. Backend API (Node.js)

The backend system handles all business logic and communication.

Features

- REST API endpoints
- Request handling and validation
- Data processing coordination
- Integration with Python automation services

14. Database Management

The system uses a relational database to store and manage data.

Data Stored

- User information
- Customer records
- Shipping requests
- Generated CSV files
- Logs and system activity

15. Admin Panel

Admin Login

Administrators can securely access the system through a dedicated login interface.

Admin Dashboard

The dashboard provides a comprehensive overview of:

- Total users
- System activity
- Generated data and files

User Management

Admins can:

- View all users
- Edit user details
- Delete users
- Monitor activity

Data Management

Admins have full access to:

- Uploaded PDF files
- Generated CSV outputs
- System logs

Analytics

- Platform usage statistics
- Data insights
- Performance tracking

16. Challenges and Solutions

Handling Unstructured PDF Data

Different PDF formats created inconsistencies in data extraction.
This was resolved using a hybrid approach combining text extraction and OCR.

Supporting Multiple Layouts

PDFs varied in structure and formatting.
Dynamic parsing logic was implemented to handle multiple layouts.

Ensuring Data Accuracy

Extracted data required validation.
Validation layers and error logging were introduced.

Automation Reliability

Automation scripts needed stability.
Retry mechanisms and logging improved reliability.

API Integration

Communication between systems required robustness.
Error handling and structured API design ensured smooth integration.

17. Outcome

The system successfully transformed manual and repetitive workflows into an automated pipeline.

Key results include:

- Significant reduction in manual data entry
- Faster processing of documents
- Improved accuracy in data handling
- Centralized management of logistics operations
- Scalable and maintainable architecture

18. Conclusion

This platform integrates logistics management, automation, and data processing into a unified system. By combining PDF data extraction, workflow automation, and structured data management, the solution enhances operational efficiency and reduces dependency on manual processes.

The architecture is designed to be scalable, allowing future enhancements such as advanced analytics, AI-based document processing, and cloud deployment.