

# CMDA Plan Permit Scraper - Requirements Document

Document Version 1.0

Date: January 2026

Prepared By: Thirumurugan Subramaniyan

| Section No. | Title                        | Subsections / Description  |
|-------------|------------------------------|--|
| 1           | Project Overview             | Project title, client, objectives, business problem, solution benefits |
| 1.1         | Project Title                | CMDA Plan Permit Scraper – Zoho CRM Integration System                 |
| 1.2         | Client                       | Ajantha Bathroom Products and Pipes Pvt. Ltd                           |
| 1.3         | Project Objective            | Automated scraping, processing, and CRM lead creation                  |
| 1.4         | Business Problem             | Manual effort, errors, delays, and tracking challenges                 |
| 1.5         | Solution Benefits            | Time saving, accuracy, scalability, reporting                          |
| 2           | Business Requirements        | Functional and non-functional requirements                             |
| 2.1         | Functional Requirements      | Scraping, extraction, processing, CRM, reporting                       |
| 2.1.1       | Data Scraping Requirements   | Website scraping, PDF download, fault tolerance                        |
| 2.1.2       | Data Extraction Requirements | Permit, applicant, architect details                                   |
| 2.1.3       | Data Processing Requirements | Validation, deduplication, sales assignment                            |
| 2.1.4       | CRM Integration Requirements | Zoho CRM OAuth, lead creation, batching                                |
| 2.1.5       | Reporting Requirements       | Excel reports, PDF summaries, statistics                               |
| 2.1.6       | Notification Requirements    | Email alerts and reports   |

| Section No. | Title                       | Subsections / Description                     |
|-------------|-----------------------------|---|
| 2.2         | Non-Functional Requirements | Performance, reliability, usability, security |
| 2.2.1       | Performance Requirements    | Processing speed, concurrency                 |
| 2.2.2       | Reliability Requirements    | Retry logic, data integrity                   |
| 2.2.3       | Usability Requirements      | GUI, progress indicators                      |
| 2.2.4       | Security Requirements       | Credential encryption, privacy                |
| 2.2.5       | Compatibility Requirements  | OS, Excel, CRM, email                         |
| 3           | User Requirements           | Roles, needs, UI expectations                 |
| 3.1         | User Roles                  | Sales Manager, Sales Team, IT Administrator   |
| 3.1.1       | Sales Manager               | Monitoring, reports, territory management     |
| 3.1.2       | Sales Team Members          | Lead follow-up, CRM updates                   |
| 3.1.3       | IT Administrator            | Configuration, troubleshooting                |
| 3.2         | User Interface Requirements | Screens, navigation, feedback                 |
| 3.2.1       | Main Screen Requirements    | Controls, progress, branding                  |
| 3.2.2       | Navigation Requirements     | Layout, consistency                           |
| 3.2.3       | Feedback Requirements       | Notifications, errors                         |
| 4           | Data Requirements           | Input and output data                         |
| 4.1         | Input Data Sources          | CMDA website, PDF documents                   |
| 4.1.1       | CMDA Website Data           | HTML tables, yearly data                      |
| 4.1.2       | PDF Documents               | Permit and architect details                  |
| 4.2         | Output Data Requirements    | Excel, CRM, reports                           |
| 4.2.1       | Excel Report Format         | Column definitions                            |

| Section No. | Title                        | Subsections / Description            |
|-------------|------------------------------|--------------------------------------|
| 4.2.2       | Zoho CRM Lead Fields         | CRM field mapping                    |
| 4.2.3       | Report Files                 | Excel, PDF, database files           |
| 5           | Process Requirements         | End-to-end workflow                  |
| 5.1         | Main Processing Workflow     | Scraping, processing, CRM, reporting |
| 5.1.1       | Scraping Process             | PDF collection and parsing           |
| 5.1.2       | Data Processing Process      | Validation, matching                 |
| 5.1.3       | CRM Integration Process      | Lead creation                        |
| 5.1.4       | Reporting Process            | Report generation                    |
| 5.2         | Business Rules               | Sales assignment, matching logic     |
| 5.2.1       | Sales Territory Assignment   | Area-based mapping                   |
| 5.2.2       | Record Matching Rules        | Matched vs unmatched                 |
| 5.2.3       | Data Validation Rules        | Field constraints                    |
| 5.2.4       | Error Handling Rules         | Retry and logging                    |
| 6           | Integration Requirements     | External system integration          |
| 6.1         | External System Integrations | Zoho CRM, Email, File system         |
| 6.1.1       | Zoho CRM Integration         | REST API, OAuth                      |
| 6.1.2       | Email System Integration     | Gmail SMTP                           |
| 6.1.3       | File System Integration      | Local storage                        |
| 6.2         | Data Exchange Requirements   | Formats and APIs                     |
| 6.2.1       | Data Formats                 | Input/output formats                 |
| 6.2.2       | API Specifications           | Endpoints and responses              |
| 7           | Quality Requirements         | Performance, security, usability     |
| 7.1         | Performance Requirements     | Speed and resource usage             |

| Section No. | Title                    | Subsections / Description          |
|-------------|--------------------------|------------------------------------|
| 7.2         | Reliability Requirements | Accuracy and recovery              |
| 7.3         | Security Requirements    | Data protection                    |
| 7.4         | Usability Requirements   | User experience                    |
| 8           | Constraints              | Technical, business, regulatory    |
| 8.1         | Technical Constraints    | OS, runtime, hardware              |
| 8.2         | Business Constraints     | CRM and process limits             |
| 8.3         | Regulatory Constraints   | Compliance and privacy             |
| 9           | Assumptions              | Technical, business, operational   |
| 9.1         | Technical Assumptions    | System stability                   |
| 9.2         | Business Assumptions     | Sales process                      |
| 9.3         | Operational Assumptions  | Daily usage                        |
| 10          | Dependencies             | External and internal dependencies |
| 10.1        | External Dependencies    | CMDA, Zoho, Gmail                  |
| 10.2        | Internal Dependencies    | Sales team, IT                     |
| 10.3        | Development Dependencies | Libraries and frameworks           |

## **1. PROJECT OVERVIEW**

### **1.1 Project Title**

CMDA Plan Permit Scraper - Zoho CRM Integration System

### **1.2 Client**

**Ajantha Bathroom Products and Pipes Pvt. Ltd**

### **1.3 Project Objective**

To develop an automated system that extracts planning permit data from the Chennai Metropolitan Development Authority (CMDA) website, processes the information, and creates leads in Zoho CRM for sales team follow-up.

### **1.4 Business Problem**

Currently, the sales team manually searches for new construction permits on the CMDA website, extracts relevant information, and manually creates leads in CRM. This process is:

- Time-consuming (4-6 hours daily)
- Prone to human error
- Inconsistent in data collection
- Delayed in lead creation
- Difficult to track and report

### **1.5 Solution Benefits**

- Time Saving: Reduce manual work from 6 hours to 15 minutes daily
- Accuracy: Ensure 100% data consistency and completeness
- Timeliness: Real-time lead creation as permits are published
- Reporting: Automated analytics and performance tracking
- Scalability: Process unlimited permits without additional resource

## **2. BUSINESS REQUIREMENTS**

### **2.1 Functional Requirements**

#### **2.1.1 Data Scraping Requirements**

- Must extract data from CMDA website for years: 2022, 2023, 2024, 2025
- Must handle different entry counts (10, 25, 50, All records per page)
- Must download PDF documents from the website
- Must extract text content from downloaded PDFs

- Must continue processing even if some records fail

### **2.1.2 Data Extraction Requirements**

- Must extract the following fields from each permit:
  - File Number
  - Planning Permission Number
  - Permit Number
  - Date of Permit
  - Date of Application
  - Mobile Number
  - Email ID
  - Applicant Name with Address
  - Nature of Development
  - Site Address
  - Area Name
- Must extract architect details from approval letters:
  - Architect Name
  - Architect Address
  - Architect Email
  - Architect Mobile Number

### **2.1.3 Data Processing Requirements**

- Must identify and extract dwelling unit information
- Must categorize records as matched or unmatched based on business rules
- Must assign sales team members based on geographical areas
- Must handle shared sales territories (e.g., "Jagan / Karthik")
- Must validate and clean extracted data
- Must compare new records with existing database to avoid duplicates

### **2.1.4 CRM Integration Requirements**

- Must authenticate with Zoho CRM using OAuth 2.0
- Must create Leads in Zoho CRM with all extracted information
- Must assign leads to correct sales team members
- Must calculate number of bathrooms based on dwelling units

- Must handle batch creation of records (up to 100 at a time)
- Must provide success/failure feedback for each record

#### **2.1.5 Reporting Requirements**

- Must generate Excel reports with all extracted data
- Must include hyperlinks to original PDF documents
- Must generate PDF summary reports with statistics
- Must track successful vs failed scrapes
- Must record new vs existing records

#### **2.1.6 Notification Requirements**

- Must send email alerts for unmatched areas
- Must send email reports for matched/unmatched records
- Must send alert when no new records are found
- Must include attachments with detailed data

### **2.2 Non-Functional Requirements**

#### **2.2.1 Performance Requirements**

- Must process 100+ PDFs within 30 minutes
- Must handle PDFs up to 10MB in size
- Must support concurrent processing of multiple years
- System response time for GUI operations: < 2 seconds

#### **2.2.2 Reliability Requirements**

- Must achieve 95% successful data extraction rate
- Must continue processing even if individual PDFs fail
- Must automatically retry failed operations
- Must maintain data integrity throughout processing

#### **2.2.3 Usability Requirements**

- Must have intuitive graphical user interface
- Must provide clear progress indicators
- Must show real-time statistics during processing
- Must have professional branding and styling
- Must require minimal training for users

#### **2.2.4 Security Requirements**

- Must securely store CRM credentials
- Must encrypt authentication tokens
- Must not store sensitive data in plain text
- Must comply with data privacy regulations

### **2.2.5 Compatibility Requirements**

- Must work on Windows 10/11 operating systems
- Must support Excel 2016 and later
- Must integrate with Zoho CRM
- Must work with Gmail for email notifications

## **3. USER REQUIREMENTS**

### **3.1 User Roles**

#### **3.1.1 Sales Manager**

- **Primary Goal:** Monitor lead generation and team performance
- **Tasks:**
  - Run the scraper for different years
  - Generate reports
  - Review unmatched areas
  - Assign territories to sales team
- **Needs:**
  - Simple interface with start/stop controls
  - Visual progress indicators
  - Comprehensive reports
  - Email notifications

#### **3.1.2 Sales Team Members**

- **Primary Goal:** Receive qualified leads in CRM
- **Tasks:**
  - Receive leads in Zoho CRM
  - Follow up with prospects
  - Update lead status
- **Needs:**



- Accurate lead information
- Timely lead delivery
- Complete contact details
- Correct territory assignment

### **3.1.3 IT Administrator**

- Primary Goal: Maintain and configure the system

#### **- Tasks:**

- Install and configure the application
- Update sales territory mappings
- Troubleshoot issues
- Manage credentials

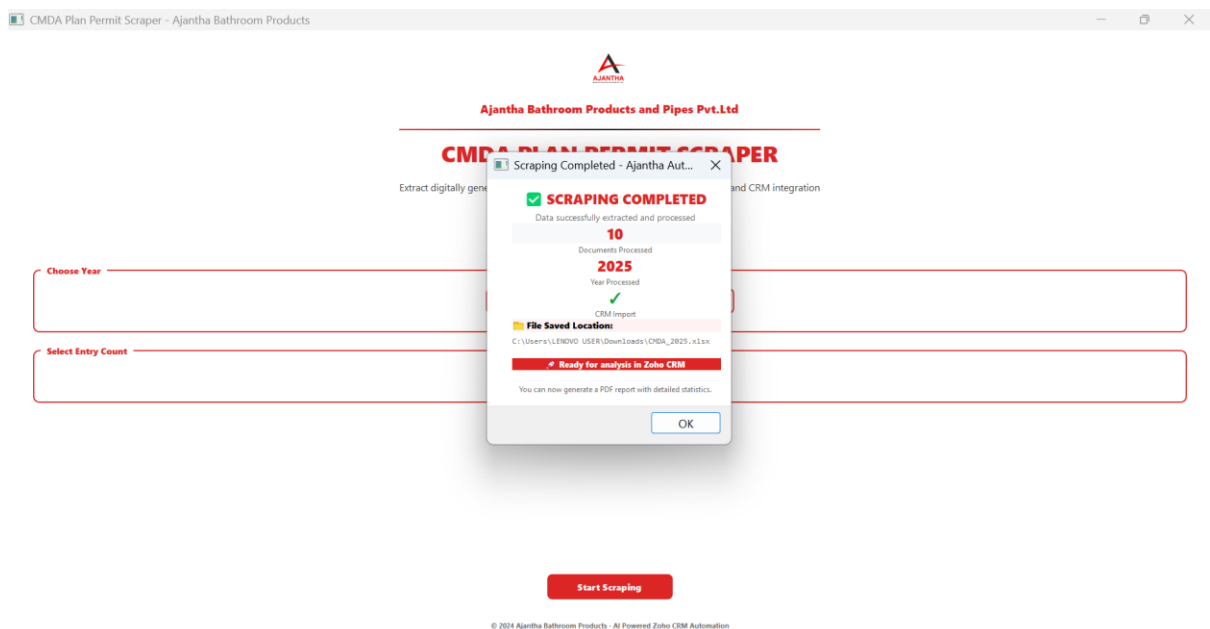
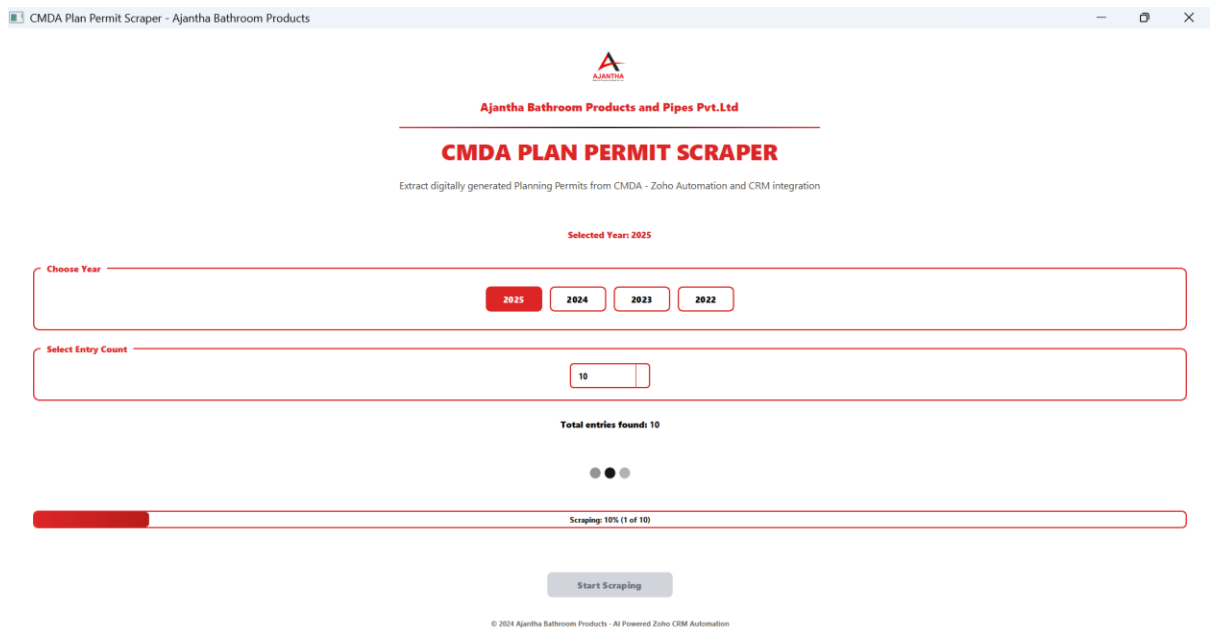
#### **- Needs:**

- Configuration files
- Error logging
- Maintenance tools
- Documentation

## **3.2 User Interface Requirements**

### **3.2.1 Main Screen Requirements**

- Display company logo and branding
- Show year selection buttons (2022-2025)
- Display entry count dropdown (10, 25, 50, All)
- Show total records found for selected year
- Include Start Scraping button
- Include Generate Report button
- Show progress bar during scraping
- Display loading animation
- Show completion message with statistics



### 3.2.2 Navigation Requirements

- Single-screen interface (no complex navigation)
- Clear visual hierarchy
- Consistent color scheme (company colors)
- Responsive layout for different screen sizes
- Tooltips for complex operations

### 3.2.3 Feedback Requirements

- Real-time progress updates
- Success/failure notifications

- File save location display
- Error messages in plain language
- Confirmation dialogs for important actions

## 4. DATA REQUIREMENTS

### 4.1 Input Data Sources

#### 4.1.1 CMDA Website Data

- Source: <https://cmdachennai.gov.in/OnlinePPAApprovalDetails/>
- Format: HTML tables with PDF links
- Frequency: Daily updates
- Years Required: 2022, 2023, 2024, 2025, 2026
- Records per Page: Variable (10, 25, 50, All)

Online PPA Approval Details Since GoLive

2025

Show 10 entries

Excel

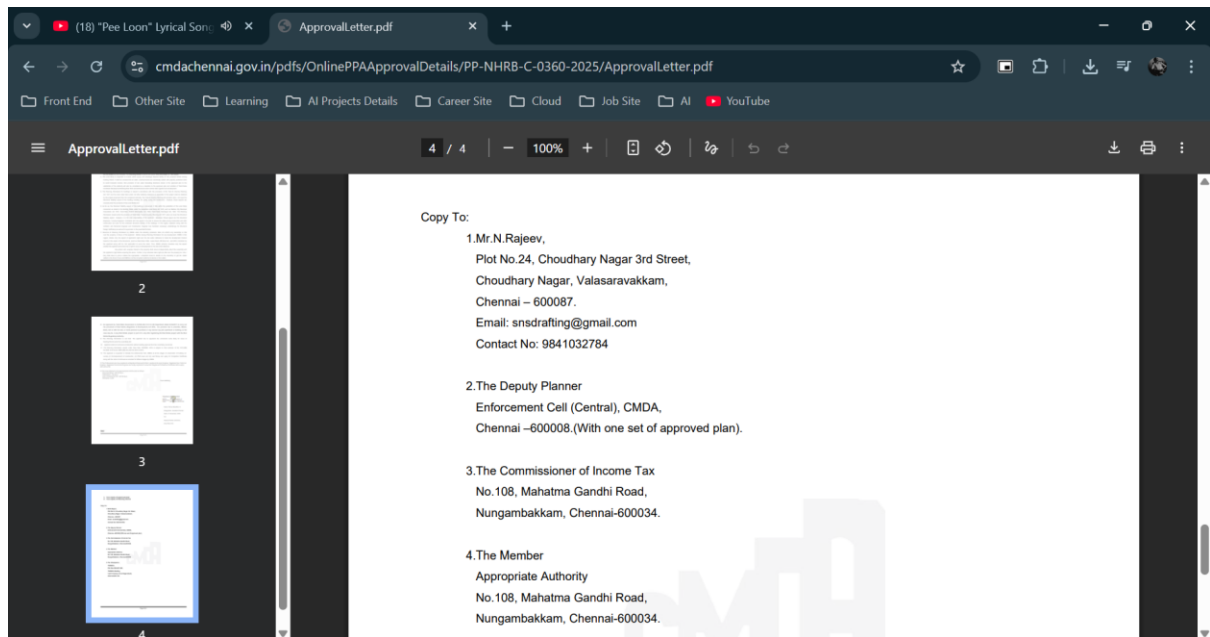
Search:

| Sl. No. | File No.                 | Permit No. | Permit Issued Date | Building Category       | Approval Letter | Approved Plan | Development Charges | Plan Per |
|---------|--------------------------|------------|--------------------|-------------------------|-----------------|---------------|---------------------|----------|
| 1       | CMDA/PP/NHRR/S/0474/2025 | OL-02509   | 31-12-2025         | Non High Rise Buildings | Approval        | Approved      | Development         | Plan     |

#### 4.1.2 PDF Documents

- Type: Planning permission documents
- Size: 100KB - 10MB per document
- Format: Scanned/Text PDFs

- Content: Permit details, applicant information, architect details



## 4.2 Output Data Requirements

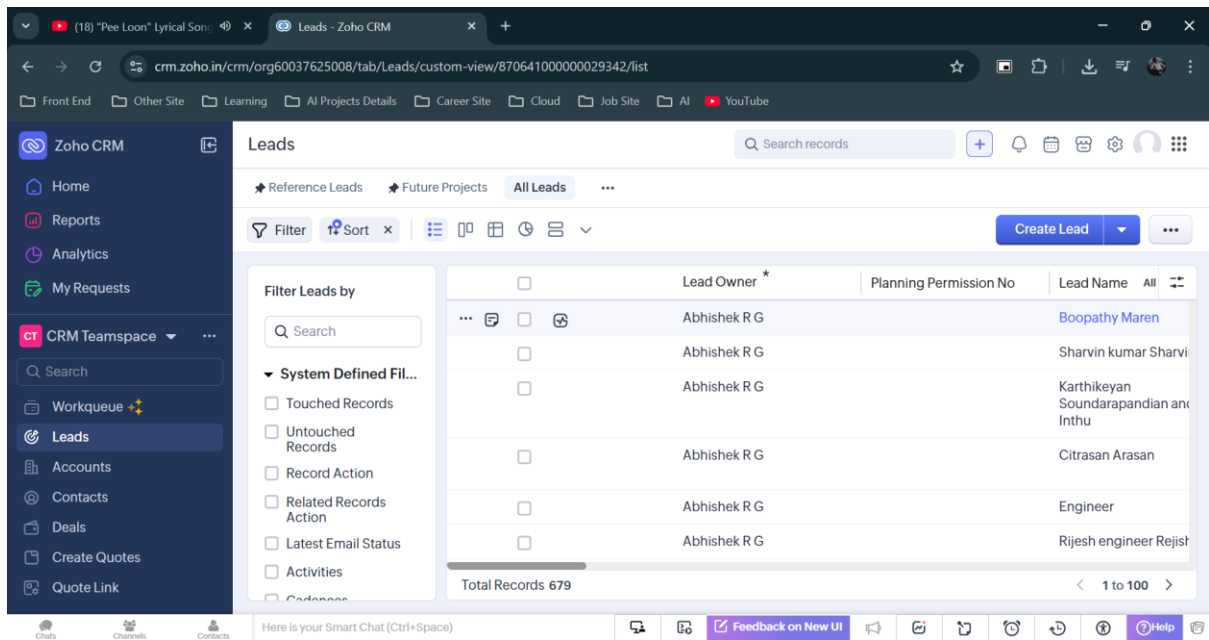
### 4.2.1 Excel Report Format

| Column Name             | Data Type | Required | Description                 |
|-------------------------|-----------|----------|-----------------------------|
| File No.                | Text      | Yes      | CMDA file number            |
| Planning Permission No. | Text      | Yes      | Planning permission number  |
| Permit No.              | Text      | Yes      | Building permit number      |
| Date of Permit          | Date      | Yes      | Permit issue date           |
| Date of Application     | Date      | Yes      | Application submission date |
| Mobile No.              | Text      | Yes      | Applicant mobile number     |
| Email ID                | Text      | Yes      | Applicant email address     |
| Applicant Name          | Text      | Yes      | Full applicant name         |
| Applicant Address       | Text      | Yes      | Complete applicant address  |
| Nature of Development   | Text      | Yes      | Type of construction        |
| Dwelling Unit Info      | Text      | Yes      | Number of units/dwellings   |
| Site Address            | Text      | Yes      | Construction site address   |
| Area Name               | Text      | Yes      | Geographical area           |

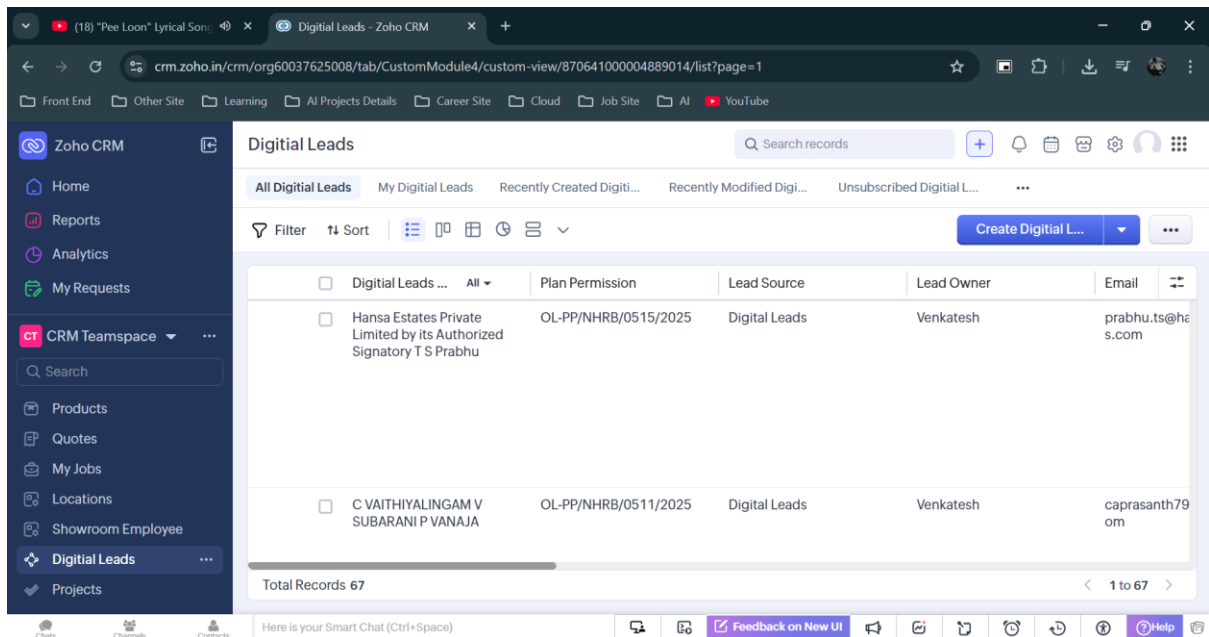
| Column Name       | Data Type | Required | Description                |
|-------------------|-----------|----------|----------------------------|
| Architect Name    | Text      | Yes      | Registered architect name  |
| Architect Address | Text      | Yes      | Architect office address   |
| Architect Email   | Text      | No       | Architect email address    |
| Architect Mobile  | Text      | No       | Architect mobile number    |
| View Online       | Hyperlink | Yes      | Link to original PDF       |
| Approved Plan     | Hyperlink | No       | Link to approved plans     |
| Approval Letter   | Hyperlink | No       | Link to approval letter    |
| Salesperson       | Text      | Yes      | Assigned sales team member |

#### 4.2.2 Zoho CRM Lead Fields

- Lead Name (First Name + Last Name)
- Company Name
- Email
- Mobile Number
- Planning Permission Number
- Nature of Development
- Area Name
- Site Address
- Architect Name
- Number of Bathrooms (calculated)
- Number of Units (extracted)
- Lead Owner (Salesperson)
- Lead Source (Digital Leads)
- Date of Permit



The screenshot shows the Zoho CRM interface with the 'Leads' module selected. The left sidebar contains navigation options like Home, Reports, Analytics, My Requests, CRM Teamspace, and a search bar. The main content area displays a list of leads under the 'All Leads' tab. A filter sidebar on the left allows filtering by system-defined filters such as 'Touched Records', 'Untouched Records', 'Record Action', 'Related Records Action', 'Latest Email Status', and 'Activities'. The lead list includes columns for checkboxes, Lead Owner, Planning Permission No, Lead Name, and an 'All' column. The first lead is 'Boopathy Maren' owned by 'Abhishek R G'. The bottom of the list shows 'Total Records 679' and a pagination control for '1 to 100'.



The screenshot shows the Zoho CRM interface with the 'Digital Leads' module selected. The left sidebar is similar to the previous screenshot but includes additional options like 'Products', 'Quotes', 'My Jobs', 'Locations', 'Showroom Employee', 'Digital Leads', and 'Projects'. The main content area displays a list of digital leads under the 'All Digital Leads' tab. The filter sidebar on the left allows filtering by system-defined filters. The lead list includes columns for checkboxes, Digital Leads, Plan Permission, Lead Source, Lead Owner, and Email. The first lead is 'Hansa Estates Private Limited by its Authorized Signatory T S Prabhu' with plan permission 'OL-PP/NHRB/0515/2025' and lead source 'Digital Leads'. The bottom of the list shows 'Total Records 67' and a pagination control for '1 to 67'.

### 4.2.3 Report Files

- Excel Report: Saved in Downloads folder with timestamp
- PDF Summary Report: Generated on demand
- Temporary Files: Stored in temp directory during processing
- Database File: "ExistData.xlsx" for tracking processed records

| File No. | Permit No. | Date of permit | Applicant Name   | Address  | Area Name                          | Architect Name | Architect Address | Architect Email | Architect Mobile | Online View | Download |
|----------|------------|----------------|--|--|------------------------------------|----------------|-------------------|-----------------|------------------|-------------|----------|
| 1        | 16-06-2022 | 19/10/2022     | 90250917 mohamed TANSIDCO 1, MSB 1, S The Planni                                       | Thiru. T.V. No.12, Jan Building : f 79 dwelling, Old Door f Mylapore | Thiru. A. N No. 4, Polir           |                |                   |                 |                  | View PDF    | View A   |
| 2        | 25-06-2022 | 22/03/2022     | 944400934 venran200 MS PRAGA 8/9, C 27, (Planning P 8 dwelling Plot No.C-: Adambakk    | comprised Adambakk Thiru. Mar Ram Niwa:                              |                                    |                |                   |                 |                  | View PDF    | View A   |
| 3        | 12-06-2022 | 18/05/2022     | 96771331 design@k Kishor Got 5, 1, Bisho The Planni                                    |  |                                    |                |                   |                 |                  | View PDF    | View A   |
| 4        | 25-03-2022 | 04/03/2022     | 98406992 sreepathy R Revathik 7 & 8, Flat Planning P 13 dwellin, Pot Nos. 7 : Adambakk |  |                                    |                |                   |                 |                  | View PDF    | View A   |
| 5        | 29-01-2022 | 11/09/2022     | 87545808 athulyame ATHULYA 5/8, 1, Thir Proposed : 11 dwelling, Door No.5 Adambakk     |  |                                    |                |                   |                 |                  | View PDF    | View A   |
| 6        | 28-07-2022 | 02/12/2022     | 94441909 gomathi@ P VISHNUC 110, 110, S Proposed i                                     | Millinium : Adayalam   |                                    |                |                   |                 |                  | View PDF    | View A   |
| 7        | 17-10-2022 | 19/06/2022     | 97419510 shreanikd ASHOK KU Tiruvottiy Planning P Single dwe Ranjith Ro Adayar         |  |                                    |                |                   |                 |                  | View PDF    | View A   |
| 8        | 08-12-2022 | 26/05/2022     | 91500201 kceeltdpur KCEE PROF 8/19, A, 2n Planning P 10 Dwelling Old Door f Adyar      |  |                                    |                |                   |                 |                  | View PDF    | View A   |
| 9        | 12-11-2022 | 06/08/2022     | 99620309 Canilreddyn Venugopal 32, na, Dh: Proposed : 5 dwelling Plot No.30 Adyar      |  |                                    |                |                   |                 |                  | View PDF    | View A   |
| 10       | 28-08-2022 | 07/01/2022     | 984104801 sdineshkui COROMAN 600032, Ac Planning P                                     | abutting G Adyar   | Thiru. Dhir Reg. No. C. dhinesh.ar |                |                   |                 |                  | View PDF    | View A   |
| 11       | 26-06-2022 | 08/01/2022     | 83000234 share2msc ABDUL KAI 349/125, N Proposed                                       | Old Nos. 2 Adyar   |                                    |                |                   |                 |                  | View PDF    | View A   |
| 12       | 11-06-2022 | 06/02/2022     | 70922453C aperiasam PERIASAM 7, NIL, MAI Proposed : 6 Dwelling Plot No.6 f Adyar       |  |                                    |                |                   |                 |                  | View PDF    | View A   |
| 13       | 12-11-2022 | 28/04/2022     | 93423339 nagalingar M Kandavi 7/49, NIL, Planning P 20 dwelling, Plot No.2 : Agaramthe |  |                                    |                |                   |                 |                  | View PDF    | View A   |
| 14       | 25-08-2022 | 29/05/2022     | 97908817 planning@ RUBY BUIL 247, B, VEL Proposed i                                    | S.No.100/: Agaramthe   |                                    |                |                   |                 |                  | View PDF    | View A   |
| 15       | 28-03-2022 | 20/01/2022     | 93423339 nagalingar M Kandavi 7/49, NA, / Proposed :                                   | S.No.236/: Agaramthe   |                                    |                |                   |                 |                  | View PDF    | View A   |
| 16       | 19-05-2022 | 30/09/2022     | 988400881 Brindhava SMK ENTE 24 A, 24 A, Planning P                                    | S.No.84/1: Agraharan   |                                    |                |                   |                 |                  | View PDF    | View A   |
| 17       | 16-10-2022 | 03/07/2022     | 994037967 sunassocie MV Prop 71, 71, Kal Planning P                                    | S.Nos. 177 Alamathi  |                                    |                |                   |                 |                  | View PDF    | View A   |

**CMDA SCRAPING REPORT**  
Ajanta Bathroom Products - 2025

Report Generated: 2025-01-08 15:23:56  
Data File: CMDA\_2025.xlsx

### 1. Scrapping Statistics

| Metric                  | Count |
|-------------------------|-------|
| Total Records Attempted | 10    |
| Successfully Scraped    | 10    |
| Failed to Scrape        | 0     |

Failed to Scrape File Numbers: None

### 2. CRM Integration Results

| CRM Import Status | Successful   |
|-------------------|--|
| Message           | No new records to process. Excel file is up to date. |
| Status Code       | 200  |

### 3. Data Analysis

New Records Count: 0

| Record Type       | Count |
|-------------------|-------|
| Matched Records   | 0     |
| Unmatched Records | 0     |

Unmatched Areas Count: 0

## 5. PROCESS REQUIREMENTS

### 5.1 Main Processing Workflow

#### 5.1.1 Scrapping Process

1. User selects year (2022-2025)
2. System fetches record count from CMDA website
3. User selects number of records to process
4. System scrapes PDF links from website

5. System downloads PDF documents
6. System extracts text from PDFs
7. System parses structured data from text
8. System extracts architect information

#### **5.1.2 Data Processing Process**

1. Validate extracted data
2. Clean and normalize text
3. Extract area names from addresses
4. Match areas to sales territories
5. Assign sales team members
6. Calculate derived fields (bathrooms, square feet)
7. Compare with existing records
8. Identify new records

#### **5.1.3 CRM Integration Process**

1. Authenticate with Zoho CRM
2. Format data for CRM API
3. Create leads in batches
4. Assign leads to sales team
5. Record success/failure status
6. Update existing records database

#### **5.1.4 Reporting Process**

1. Generate Excel report with all data
2. Create PDF summary report
3. Send email notifications
4. Save reports to designated folders
5. Clean up temporary files

### **5.2 Business Rules**

#### **5.2.1 Sales Territory Assignment**

SALES TEAM TERRITORIES:

- Abhishek R G: Adambakkam, Alandur, Guindy, Madipakkam, etc.
- Jagan: Adyar, Egmore, Koyambedu, Mogappair, etc.



- Karthik: Arumbakkam, Kodambakkam, Valasaravakkam, etc.
- Venkatesh: Anakaputhur, Kundrathur, Pallavaram, Tambaram, etc.
- Dinakaran: Kottivakkam, Sholinganallur, Thiruvananthapuram, etc.
- Balachander: Maduravoyal, Ayapakkam, etc.
- Shared Territories: Jagan/Balachander, Karthik/Venkatesh, etc.

### **5.2.2 Record Matching Rules**

- Matched Records:
  - Have "Dwelling Unit Info" populated, OR
  - Nature of Development contains keywords: school, hospital, college, etc.
- Unmatched Records: Do not meet above criteria
- New Records: Not found in existing database

### **5.2.3 Data Validation Rules**

- Mobile numbers must be 10 digits
- Email addresses must contain "@" and "."
- Dates must be in valid format
- Required fields cannot be empty
- File numbers must start with "CMDA"

### **5.2.4 Error Handling Rules**

- Continue processing if single PDF fails
- Mark failed records in report
- Send alert for critical errors
- Retry failed operations once
- Log all errors for troubleshooting

## **6. INTEGRATION REQUIREMENTS**

### **6.1 External System Integrations**

#### **6.1.1 Zoho CRM Integration**

- Interface: REST API
- Authentication: OAuth 2.0
- Data Format: JSON
- Rate Limits: 100 records per batch
- Required Modules: Leads, Users

- Fields to Update: All lead fields as specified
- Error Handling: Retry on failure, log errors

### **6.1.2 Email System Integration**

- Service: Gmail SMTP
- Protocol: SMTP over SSL
- Port: 465
- Authentication: App-specific password
- Attachments: Excel files up to 25MB
- Format: HTML emails with styling
- Recipients: Configurable via .env file

### **6.1.3 File System Integration**

- Input: PDF documents from web
- Output: Excel files, PDF reports
- Storage: Local file system
- Locations: Downloads folder, Temp folder
- Permissions: Read/Write access required

## **6.2 Data Exchange Requirements**

### **6.2.1 Data Formats**

- Input: PDF, HTML, JSON
- Output: Excel (.xlsx), PDF, JSON
- Intermediate: Python objects, Data Frames
- Configuration: .env files, JSON files

### **6.2.2 API Specifications**

- Zoho CRM API Version: v2
- HTTP Methods: GET, POST
- Headers: Authorization, Content-Type
- Response Codes: Handle 200, 201, 400, 401, 429
- Pagination: Not required (batch processing)

## **7. QUALITY REQUIREMENTS**

### **7.1 Performance Requirements**

- Scraping Speed: Process 100 records in  $\leq 30$  minutes

- PDF Processing: Extract text in  $\leq 10$  seconds per PDF
- CRM Integration: Create 100 leads in  $\leq 2$  minutes
- Memory Usage:  $\leq 500$ MB during peak operation
- CPU Usage:  $\leq 50\%$  on average system

## **7.2 Reliability Requirements**

- Availability: 99% during business hours
- Success Rate: 95% of records processed successfully
- Data Accuracy: 98% field extraction accuracy
- Error Recovery: Automatic recovery from network failures
- Data Integrity: No data loss during processing

## **7.3 Security Requirements**

- Authentication: Secure OAuth 2.0 flow
- Credentials: Encrypted storage in .env file
- Data Protection: No sensitive data in logs
- Access Control: Local system only, no network exposure
- Compliance: Basic GDPR principles for data handling

## **7.4 Usability Requirements**

- Learnability: Users proficient in 30 minutes
- Efficiency: Complete scraping in 3 clicks
- Error Prevention: Clear warnings before destructive actions
- Satisfaction: Professional interface meeting corporate standards
- Accessibility: Clear text, adequate contrast, simple navigation

# **8. CONSTRAINTS**

## **8.1 Technical Constraints**

- Must run on Windows 10/11 only
- Requires Chrome/Chromium browser installation
- Requires Python 3.8+ runtime
- Requires 4GB RAM minimum
- Requires stable internet connection
- Requires Zoho CRM account with API access

## **8.2 Business Constraints**

- Cannot modify CMDA website data
- Must respect Zoho CRM API rate limits
- Must comply with data privacy regulations
- Must maintain sales territory assignments
- Must preserve existing CRM data structure
- Must not disrupt current sales processes

## **8.3 Regulatory Constraints**

- Must not scrape data outside permitted use
- Must respect website terms of service
- Must handle personal data responsibly
- Must provide opt-out mechanism if required
- Must maintain data processing records

# **9. ASSUMPTIONS**

## **9.1 Technical Assumptions**

- CMDA website structure remains stable
- PDF document format remains consistent
- Zoho CRM API remains available
- Internet connectivity is reliable
- User has administrative rights on computer
- Antivirus software won't block the application

## **9.2 Business Assumptions**

- Sales territory mappings are correct
- All sales team members use Zoho CRM
- Email notifications are desired
- Reports need to be saved locally
- Historical data from 2022-2025 is needed
- Process will run daily during business hours

## **9.3 Operational Assumptions**

- IT department will handle installation
- Sales manager will run the application

- Email alerts will be monitored
- Reports will be reviewed regularly
- System will be updated as needed
- Backup of data will be maintained

## **10. DEPENDENCIES**

### **10.1 External Dependencies**

- CMDA website availability and structure
- Zoho CRM API availability and stability
- Gmail SMTP service availability
- Chrome/Chromium browser compatibility
- Python package availability and updates

### **10.2 Internal Dependencies**

- Sales team territory assignments
- CRM user account permissions
- Email account credentials
- Network firewall configurations
- IT support for installation and maintenance

### **10.3 Development Dependencies**

- Playwright browser automation library
- PDF processing libraries (pdfplumber, PyPDF2)
- Data processing libraries (pandas, openpyxl)
- GUI framework (PyQt5)
- Report generation libraries (reportlab)