Product Requirement Document (PRD)

Field	Value
Project	Evergreen AI
Client	Evergreen Group Pte Ltd
Contact	Mrs Wee
Phone	+65 9683 1343 / +65 9424 4633
Email	mrswee@evergreen.com.sg
Address	8 New Industrial Road, #01-02/03, LHK3, Singapore 536200
Version	1.1
Status	Draft
Author	Aira Ling
Reviewers	Ken Ling
Created	2026-02-24
Last Updated	2026-02-24

Revision History

Version	Date	Author	Changes
1.0	2026-02-24	Aira Ling	Initial PRD
1.1	2026-02-24	Aira Ling	Added Quotation Generation epic (FR-019 to FR-022), client details

1. Executive Summary

Evergreen AI is a web-based AI assistant for Evergreen Group Pte Ltd, a Singapore office supplies company (evergreen.com.sg) serving corporates and schools. The system addresses three core operational bottlenecks: manual document processing for sales operations, labour-intensive e-commerce content creation for marketplace listings, and time-consuming quotation preparation.

Workstream 1 — Document Processing: Incoming invoices, purchase orders, and delivery orders (from email, WhatsApp, and local folders) are OCR-scanned and AI-parsed into structured data. A review UI lets staff verify extractions side-by-side with originals, then export to Excel for ERP import. Extracted line items are matched against Evergreen's product catalogue.

Workstream 2 — E-commerce Marketing Agent: AI generates optimised product listing content for Shopee.sg and Lazada.sg — including enhanced product photos (background removal, lifestyle shots) and platform-specific titles, descriptions, and bullet points. A review queue ensures human approval before content goes live.

Workstream 3 — Quotation Generation: AI pulls customer sales history from Evergreen's existing server, analyses purchase patterns, and generates smart draft quotations pre-populated with likely products and historical pricing. Sales staff review, adjust, and export professional branded PDFs — or email quotations directly to customers.

All three workstreams share a common product catalogue synced via periodic CSV/Excel imports from Evergreen's ERP (no API). The system targets 4 users (2 sales, 2 marketing) with a 4-week timeline and SGD 50K budget.

2. Problem Statement & Hypothesis

Problem

Evergreen's sales team manually processes 30–50+ supplier/customer documents daily — invoices, POs, and delivery orders arriving via email, WhatsApp, and paper. Each document requires 5–10 minutes of manual data entry into spreadsheets for ERP import. This is error-prone, tedious, and consumes ~4 hours/day across 2 staff. Meanwhile, the marketing team manually creates product listings for Shopee and Lazada — photographing products, removing backgrounds, writing descriptions — spending 20–30 minutes per SKU. With 500+ products to list, this backlog grows faster than the team can process.

Hypothesis

If we build an AI-powered document processing pipeline with human review, then the sales team will reduce per-document processing time by 70%+ (from 5–10 min to 1–2 min verification). If we build an AI marketing content generator with review queue, then the marketing team will increase listing output by 5x (from ~15 SKUs/day to ~75 SKUs/day) while maintaining quality through human approval.

3. User Personas

Persona: Sales / Document Processor

Attribute	Details
Who	Sales admin staff handling document intake. Comfortable with Excel and basic web apps. Age 30–50.
Goals	Process incoming documents quickly and accurately, get data into ERP without errors
Frustrations	Repetitive data entry, misreading handwritten/scanned documents, switching between email/WhatsApp/folders, manual product code matching
Frequency	Daily, throughout the workday
Devices	Desktop (primary), occasionally mobile for WhatsApp documents

Persona: Marketing / E-commerce Specialist

Attribute	Details
Who	Marketing staff managing Shopee and Lazada listings. Familiar with marketplace platforms and basic image editing. Age 25–40.
Goals	Create high-quality product listings quickly, optimise for marketplace search/conversion

Frustrations	Tedious photo editing, writing repetitive descriptions, different format requirements per platform, slow listing throughput
Frequency	Daily, batch processing products
Devices	Desktop (primary)

Persona: Operations Manager

Attribute	Details
Who	Oversees both sales and marketing operations. Needs visibility into processing volumes and bottlenecks.
Goals	Monitor team throughput, ensure data quality, track catalogue sync status
Frustrations	No visibility into processing pipeline, manual status tracking
Frequency	Daily check-ins, weekly reviews
Devices	Desktop

4. User Journey Maps

Journey: Document Processing (Sales)

```
Document arrives (email/WhatsApp/folder)
    → System detects and queues document
    → OCR + AI extraction runs automatically
    → Document appears in Review Queue with status "Pending Review"
    → Staff clicks document → Side-by-side view (original | extracted data)
    → Staff verifies/corrects extracted fields
    → Staff marks as "Verified"
    → Staff selects verified documents → Export to Excel (ERP format)
    → Staff imports Excel into ERP
```

Happy path: Document is clearly scanned, AI extracts all fields correctly, staff confirms with minimal edits, exports to Excel. **Drop-off risks:** Poor scan quality → low extraction confidence → more manual corrections needed. Unrecognised document format → AI can't parse → manual entry fallback.

Journey: E-commerce Content Generation (Marketing)

```
Staff selects product(s) from catalogue
    → AI generates content: enhanced photos, title, description, bullet points
    → Content appears in Review Queue per platform (Shopee / Lazada)
    → Staff reviews AI-generated content
    → Staff edits/approves each element (photo, title, description)
    → Staff copies approved content to marketplace platform (Phase 1: manual)
    → Staff marks listing as "Published"
```

Happy path: Product has good source photo and complete catalogue data, AI generates compelling content, staff approves with minor tweaks. **Drop-off risks:** Poor source photo → AI enhancement inadequate →

manual photo editing needed. Incomplete catalogue data → AI generates generic content → extensive editing required.

Journey: Product Catalogue Sync

```
Admin exports product list from ERP as CSV/Excel
    → Admin uploads file to Evergreen AI web app
    → System parses and validates data
    → System shows diff: new products, updated products, removed products
    → Admin confirms import
    → Catalogue updated for both workstreams
```

5. Feature Requirements (Functional)

Epic: Document Ingestion

FR-001: Email Document Ingestion

Field	Value
Priority	Must
Epic	Document Ingestion
Description	System monitors a configured email inbox (IMAP/API) for incoming messages with PDF/image attachments. Attachments are automatically downloaded, queued for OCR processing, and linked to the sender's email address. Supported formats: PDF, JPEG, PNG, TIFF.

Acceptance Criteria:

1. **Given** a new email arrives with a PDF attachment, **When** the system polls the inbox, **Then** the attachment is downloaded, a document record is created with status "queued", and the sender's email is recorded.
2. **Given** an email has multiple attachments, **When** the system processes it, **Then** each attachment is created as a separate document record, all linked to the same source email.
3. **Given** an email has no attachments (text only), **When** the system processes it, **Then** it is ignored (no document created).
4. **Given** a duplicate attachment (same filename + hash from same sender within 24 hours), **When** detected, **Then** the system skips it and logs a duplicate warning.

Notes: Polling frequency: every 5 minutes (configurable). Initial setup requires client to provide email credentials or forward documents to a dedicated inbox.

FR-002: WhatsApp Document Ingestion

Field	Value
Priority	Must
Epic	Document Ingestion

Description	System receives documents sent via WhatsApp (using WhatsApp Business API or GreenAPI). Staff can forward supplier/customer documents to a dedicated WhatsApp number. Photos and PDFs are captured and queued for OCR.

Acceptance Criteria:

1. **Given** a user sends a photo of a document via WhatsApp, **When** the webhook receives the message, **Then** the image is downloaded, a document record is created with status "queued", and the sender's phone number is recorded.
2. **Given** a user sends a PDF via WhatsApp, **When** the webhook receives it, **Then** the PDF is processed identically to email PDFs.
3. **Given** a user sends a text message (no media), **When** the webhook receives it, **Then** it is acknowledged with a reply: "Please send documents as photos or PDFs."
4. **Given** multiple images sent in quick succession (within 2 minutes from same sender), **When** processed, **Then** they are grouped as pages of a single multi-page document.

Notes: Use GreenAPI for WhatsApp integration (same as PilotNow). The dedicated WhatsApp number is shared with clients/suppliers for document forwarding.

FR-003: Local Folder Ingestion

Field	Value
Priority	Must
Epic	Document Ingestion
Description	System monitors a configured local/network folder for new PDF and image files. Files dropped into the folder are automatically picked up and queued for OCR processing. Processed files are moved to a "processed" subfolder.

Acceptance Criteria:

1. **Given** a new PDF is placed in the watched folder, **When** the file watcher detects it, **Then** the file is copied to system storage, a document record is created with status "queued", and the original is moved to the "processed" subfolder.
2. **Given** a non-supported file type is placed in the folder (e.g., .docx), **When** detected, **Then** it is moved to an "unsupported" subfolder and logged.
3. **Given** the watched folder is not accessible (network drive offline), **When** the watcher fails, **Then** an error is logged and admin is notified; the system retries every 5 minutes.

Notes: This supports the client's existing workflow of saving scanned documents to a shared drive.

Epic: OCR & Data Extraction

FR-004: OCR Processing

Field	Value
Priority	Must
Epic	OCR & Data Extraction

	Documents in the queue are processed through OCR to extract raw text. Supports
Description	printed text in English and Chinese. Handles scanned PDFs, photos of documents (including skewed/rotated), and native-text PDFs (extract text directly without OCR).

Acceptance Criteria:

1. **Given** a queued document (scanned PDF or image), **When** OCR runs, **Then** raw text is extracted with 95%+ accuracy for clearly printed documents.
2. **Given** a native-text PDF (selectable text), **When** processed, **Then** text is extracted directly without OCR (faster, 100% accurate).
3. **Given** a skewed or rotated document image, **When** OCR runs, **Then** the system auto-corrects orientation before extraction.
4. **Given** OCR completes, **When** the text is extracted, **Then** the document status changes to "extracted" and the raw text is stored.
5. **Given** OCR fails (unreadable document), **When** the error occurs, **Then** the document status changes to "extraction_failed" and it appears in the review queue with a warning.

Notes: Use Google Cloud Vision API or AWS Textract for OCR. Support for handwritten text is Phase 2.

FR-005: AI Data Extraction (Structured Parsing)

Field	Value
Priority	Must
Epic	OCR & Data Extraction
Description	After OCR, an LLM parses the raw text into structured fields based on the document type (invoice, PO, delivery order). The AI identifies: document type, document number, date, company name, address, line items (product name, quantity, unit price, total, product code/SKU), subtotal, tax, grand total, payment terms, and any other relevant fields.

Acceptance Criteria:

1. **Given** OCR text from an invoice, **When** the LLM processes it, **Then** it extracts: document type ("invoice"), invoice number, date, supplier/customer name, line items (with product name, qty, unit price, line total), subtotal, GST, grand total.
2. **Given** OCR text from a purchase order, **When** the LLM processes it, **Then** it extracts: document type ("PO"), PO number, date, buyer/supplier name, line items, delivery date, payment terms.
3. **Given** OCR text from a delivery order, **When** the LLM processes it, **Then** it extracts: document type ("DO"), DO number, date, sender/recipient, line items (product name, qty — no pricing expected).
4. **Given** the LLM cannot confidently extract a field, **When** the result is returned, **Then** that field is marked with a confidence score < 0.7 and highlighted in the review UI for manual verification.
5. **Given** extraction completes, **When** the structured data is stored, **Then** the document status changes to "pending_review".

Notes: Use Gemini or GPT-4o for extraction. Provide few-shot examples of each document type in the prompt. Confidence scoring helps prioritise review effort.

FR-006: Product Catalogue Matching

Field	Value

Priority	Must
Epic	OCR & Data Extraction
Description	After data extraction, each line item is matched against Evergreen's product catalogue. Matching uses product name fuzzy matching, SKU/product code exact matching, and AI-assisted semantic matching. Matched products show the internal product code; unmatched items are flagged for manual matching.

Acceptance Criteria:

1. **Given** an extracted line item with SKU "EG-A4-80", **When** catalogue matching runs, **Then** it finds an exact match in the catalogue and links the item to the internal product record.
2. **Given** an extracted line item with product name "A4 Copy Paper 80gsm" but no SKU, **When** catalogue matching runs, **Then** it performs fuzzy name matching and returns top 3 candidates with confidence scores.
3. **Given** no match is found (confidence < 0.5 for all candidates), **When** the result is returned, **Then** the line item is flagged as "unmatched" in the review UI, and staff can manually select from the catalogue or mark as new product.
4. **Given** a staff member manually matches a product, **When** they save the match, **Then** the system learns this mapping for future documents from the same supplier (supplier-specific product name → internal SKU mapping).

Notes: Build a mapping table over time: supplier product names → Evergreen product codes. This improves accuracy as more documents are processed.

Epic: Document Review

FR-007: Document Queue / List View

Field	Value
Priority	Must
Epic	Document Review
Description	Web-based table listing all processed documents. Columns: document type, document number, date, company name, total amount, source (email/WhatsApp/folder), status (pending_review/verified/exported/failed), received date, confidence score. Filterable by status, type, source, date range. Sortable by any column.

Acceptance Criteria:

1. **Given** a user opens the Document Queue page, **When** the page loads, **Then** it shows a paginated table of all documents, most recent first, with key columns visible.
2. **Given** documents exist with various statuses, **When** user filters by "pending_review", **Then** only unverified documents are shown.
3. **Given** a document row, **When** the user clicks it, **Then** the side-by-side review view opens (FR-008).
4. **Given** the table has 500+ documents, **When** scrolling/paginating, **Then** performance remains smooth (virtual scrolling or server-side pagination).
5. **Given** a user wants to bulk export, **When** they select multiple verified documents and click "Export", **Then** a single Excel file is generated containing all selected documents.

FR-008: Side-by-Side Review View

Field	Value
Priority	Must
Epic	Document Review
Description	When a user clicks a document in the queue, a split-screen view opens. Left panel: original scanned document image (zoomable, scrollable for multi-page). Right panel: extracted data in an editable form/table. Fields with low confidence are highlighted in yellow. User can edit any field, then mark the document as "verified".

Acceptance Criteria:

1. **Given** a user opens a document in review, **When** the view loads, **Then** the left panel shows the original document image and the right panel shows all extracted fields in editable form inputs.
2. **Given** a field has confidence < 0.7, **When** displayed, **Then** it is highlighted in yellow with a tooltip showing the confidence score.
3. **Given** a user edits a field value, **When** they modify the text, **Then** the change is saved to the extracted data (auto-save or explicit save button).
4. **Given** line items are extracted, **When** displayed, **Then** they appear in an editable table with columns: product name, matched product code, quantity, unit price, line total. Users can add/remove rows.
5. **Given** a user has reviewed and corrected all fields, **When** they click "Verify", **Then** the document status changes to "verified" and it moves out of the pending queue.
6. **Given** a multi-page document, **When** the user views it, **Then** they can navigate between pages in the left panel while the extracted data remains on the right.

Notes: Consider highlighting the corresponding area in the original document when a field is focused (visual linking). This is a nice-to-have for Phase 1.

FR-009: Excel Export (ERP Format)

Field	Value
Priority	Must
Epic	Document Review
Description	Verified documents can be exported to an Excel file in Evergreen's specific ERP import format. The format includes specific column headers, data ordering, and sheet structure as defined by the client. Single or bulk export supported.

Acceptance Criteria:

1. **Given** one or more verified documents are selected, **When** user clicks "Export to Excel", **Then** an .xlsx file is downloaded with data in the client's ERP format.
2. **Given** the client's ERP format requires specific columns (e.g., "Doc No", "Doc Date", "Item Code", "Description", "Qty", "Unit Price", "Amount"), **When** the Excel is generated, **Then** columns match exactly including headers, order, and data types.
3. **Given** an exported document, **When** the export completes, **Then** the document status changes to "exported" and the export date is recorded.
4. **Given** a line item has no matched product code, **When** exported, **Then** the item code field is left blank (staff fills in manually in ERP) or uses the closest match if available.

5. **Given** the ERP format changes, **When** the admin updates the format configuration, **Then** future exports use the new format.

Notes: ERP format template will be provided by the client during onboarding. Build a configurable template engine so format changes don't require code changes.

Epic: Product Catalogue

FR-010: Catalogue CSV/Excel Import

Field	Value
Priority	Must
Epic	Product Catalogue
Description	Admin uploads a CSV or Excel file containing the product catalogue exported from their ERP. The system parses it, validates the data, shows a preview with diff (new/updated/removed products), and imports upon confirmation. This is the primary mechanism for keeping the catalogue in sync.

Acceptance Criteria:

1. **Given** an admin uploads a CSV/Excel file, **When** the system parses it, **Then** it shows a summary: X new products, Y updated, Z removed (compared to current catalogue).
2. **Given** the preview is shown, **When** the admin confirms, **Then** the catalogue is updated — new products added, existing products updated, removed products marked inactive (not deleted).
3. **Given** the file has validation errors (missing required fields, invalid data), **When** parsed, **Then** errors are shown per row and the admin can fix and re-upload.
4. **Given** a successful import, **When** complete, **Then** the import timestamp and stats are logged, and both workstreams immediately see updated catalogue data.

Notes: Required fields: product code/SKU, product name, category, unit price. Optional: description, brand, UOM, image URL. Mapping of CSV columns to system fields is configurable.

FR-011: Catalogue Browse & Search

Field	Value
Priority	Must
Epic	Product Catalogue
Description	Web-based interface to browse and search the product catalogue. Supports full-text search by name, SKU, category, and brand. Used by both sales (for manual product matching) and marketing (for selecting products to generate content).

Acceptance Criteria:

1. **Given** a user opens the catalogue page, **When** it loads, **Then** products are displayed in a paginated table/grid with key fields: SKU, name, category, brand, unit price, image.
2. **Given** a user types in the search bar, **When** searching for "A4 paper", **Then** results include all products with "A4" and "paper" in name, description, or category (full-text search).

3. **Given** search results, **When** a user clicks a product, **Then** a detail view shows all product information including any existing marketplace content.

Epic: E-commerce Content Generation

FR-012: Product Photo Enhancement

Field	Value
Priority	Must
Epic	E-commerce Content Generation
Description	AI-powered image processing for product photos. Features: background removal (replace with white/custom background), image quality enhancement (brightness, contrast, sharpness), and lifestyle/context shot generation (AI places product in a relevant setting — e.g., pen on desk, stationery in office). Input: raw product photo. Output: multiple variants for marketplace listing.

Acceptance Criteria:

1. **Given** a user uploads a raw product photo, **When** background removal runs, **Then** the product is isolated and placed on a clean white background, output as a PNG with transparent background option.
2. **Given** a product photo, **When** enhancement runs, **Then** brightness, contrast, and sharpness are auto-adjusted to produce a professional-looking product image.
3. **Given** a product and its category, **When** lifestyle shot generation runs, **Then** AI generates 2–3 contextual images (e.g., stationery on a desk, paper in a printer, pens in a pencil holder).
4. **Given** generated images, **When** displayed in the review queue, **Then** the user can approve, reject, or request regeneration of each variant.
5. **Given** a product with no uploaded photo, **When** the user tries to generate content, **Then** the system prompts to upload a photo first (or uses catalogue image if available).

Notes: Use a combination of: rembg/remove.bg API for background removal, image processing library for enhancement, and DALL-E/Stable Diffusion for lifestyle shots. Lifestyle shots are experimental — start with background removal and enhancement as core.

FR-013: Product Listing Content Generation

Field	Value
Priority	Must
Epic	E-commerce Content Generation
Description	AI generates marketplace-optimised product listing text: title, description, bullet points/key features, and SEO keywords. Content is customised per platform (Shopee.sg vs Lazada.sg) based on each platform's best practices, character limits, and style guidelines. Input: product catalogue data + photos. Output: platform-specific content sets.

Acceptance Criteria:

1. **Given** a product is selected for content generation, **When** the user chooses "Shopee", **Then** the AI generates: title (max 100 chars, keyword-optimised), description (formatted for Shopee), 5 bullet points, and suggested category path.
2. **Given** the same product, **When** the user chooses "Lazada", **Then** the AI generates: title (Lazada format), long description (HTML-formatted for Lazada), short description, and key features — different from Shopee output.
3. **Given** generated content, **When** displayed in the review UI, **Then** each element (title, description, bullets) is individually editable.
4. **Given** a product with limited catalogue data (name and price only), **When** content is generated, **Then** the AI still produces reasonable content but flags: "Limited source data — review carefully."
5. **Given** a user approves the content, **When** they click "Approve", **Then** the content is saved as the final version with status "approved" and ready for manual posting.

Notes: Train/prompt the AI with Evergreen's brand voice, competitive listings, and platform-specific guidelines. Include a "Regenerate" button with optional user instructions (e.g., "make it more professional", "emphasise eco-friendly").

FR-014: Content Review Queue

Field	Value
Priority	Must
Epic	E-commerce Content Generation
Description	Web-based review queue showing all AI-generated content awaiting approval. Each item shows: product info, generated photos, generated text (per platform), status (draft/approved/published). Marketing team reviews, edits, and approves before manual posting to marketplaces.

Acceptance Criteria:

1. **Given** content has been generated for products, **When** a marketing user opens the review queue, **Then** they see a list of products with generated content, filterable by status and platform.
2. **Given** a user clicks a product in the queue, **When** the detail view opens, **Then** they see: product photos (original + AI variants), generated text per platform, and action buttons (Approve, Edit, Regenerate, Reject).
3. **Given** a user approves content for Shopee, **When** they click Approve, **Then** the Shopee content is marked "approved" and the user can copy-paste it from a formatted view.
4. **Given** approved content, **When** the user clicks "Copy to Clipboard" for any field, **Then** the text is copied in the correct format for pasting into the marketplace platform.
5. **Given** a user marks content as "Published", **When** they confirm and optionally enter the marketplace listing URL, **Then** the status updates and the listing URL is stored for reference.

Notes: Phase 1 is manual copy-paste to Shopee/Lazada. Phase 2 would add API integration for direct publishing. The review queue should make copy-paste as frictionless as possible.

Epic: User Management & System

FR-015: User Authentication

Field	Value

Priority	Must
Epic	User Management & System
Description	Simple email + password authentication with role-based access. Roles: Admin (full access), Sales (document processing features), Marketing (e-commerce features). All roles can view the catalogue.

Acceptance Criteria:

1. **Given** a registered user, **When** they enter correct email and password, **Then** they are logged in and see features relevant to their role.
2. **Given** a Sales user logs in, **When** the dashboard loads, **Then** they see Document Queue, Review View, and Catalogue — but not E-commerce features.
3. **Given** a Marketing user logs in, **When** the dashboard loads, **Then** they see Content Generation, Review Queue, and Catalogue — but not Document Processing features.
4. **Given** an Admin user, **When** logged in, **Then** they have access to all features plus user management and system settings.

Notes: 4 users total — no need for self-registration. Admin creates accounts. Consider adding Google OAuth as a convenience in Phase 2.

FR-016: Dashboard

Field	Value
Priority	Should
Epic	User Management & System
Description	Landing page after login showing role-relevant metrics. Sales dashboard: documents processed today, pending review count, export count. Marketing dashboard: content generated today, pending approval count, published count. Admin sees both.

Acceptance Criteria:

1. **Given** a Sales user logs in, **When** the dashboard loads, **Then** they see: documents pending review (count), documents processed today, documents exported this week, and recent activity feed.
2. **Given** a Marketing user logs in, **When** the dashboard loads, **Then** they see: content pending approval (count), listings generated today, listings published this week, and recent activity feed.
3. **Given** an Admin user logs in, **When** the dashboard loads, **Then** they see combined metrics from both workstreams plus catalogue sync status.

FR-017: Activity Log / Audit Trail

Field	Value
Priority	Should
Epic	User Management & System
Description	System logs all significant actions: document processing events, review/verification actions, exports, content generation, approvals. Viewable by Admin for audit and troubleshooting.

Acceptance Criteria:

1. **Given** any user action (verify document, approve content, export, etc.), **When** the action is performed, **Then** a log entry is created with: user, action, target entity, timestamp, and details.
2. **Given** an Admin opens the activity log, **When** it loads, **Then** they can filter by user, action type, date range, and entity type.

FR-018: Notification System

Field	Value
Priority	Could
Epic	User Management & System
Description	In-app notifications for key events: new documents ready for review, content generation complete, export failures, catalogue sync complete. Optional email notifications for critical events.

Acceptance Criteria:

1. **Given** a new batch of documents finishes OCR processing, **When** they enter "pending_review", **Then** Sales users see a notification badge on the Document Queue.
2. **Given** content generation completes for a batch, **When** content is ready, **Then** Marketing users see a notification badge on the Review Queue.
3. **Given** an export or processing failure, **When** the error occurs, **Then** the relevant user is notified in-app with error details.

Epic: Quotation Generation

FR-019: Sales History Retrieval

Field	Value
Priority	Must
Epic	Quotation Generation
Description	Connect to Evergreen's existing server/database to pull customer sales history. Retrieve past transactions by customer name or company — including products purchased, quantities, unit prices, dates, and frequency. Data is used to pre-populate quotations with relevant products and historical pricing. If direct DB access is unavailable, support CSV/Excel import of sales history as fallback.

Acceptance Criteria:

1. **Given** a Sales user starts a new quotation, **When** they search/select a customer, **Then** the system retrieves that customer's purchase history showing: products, quantities, prices, and last purchase date.
2. **Given** a customer has multiple past orders, **When** history is displayed, **Then** it is sorted by most recent and shows frequency (e.g., "ordered 5 times in last 6 months").
3. **Given** the connection to the existing server is unavailable, **When** a user tries to pull history, **Then** the system shows a clear error and offers manual entry as fallback.

4. **Given** sales history is imported via CSV/Excel, **When** the file is uploaded, **Then** the system parses and indexes it by customer for quotation lookup.

Notes: Need to work with Mrs Wee's team to determine data access method — direct DB connection, API, or periodic export. Prioritise whatever their IT team can support.

FR-020: AI-Assisted Quotation Builder

Field	Value
Priority	Must
Epic	Quotation Generation
Description	AI generates a draft quotation based on customer sales history. The system analyses past purchases to suggest likely products, quantities (based on order patterns), and pricing (using last sold price or current catalogue price). Sales user can add/remove/edit line items, adjust quantities and pricing, apply discounts, and set validity period.

Acceptance Criteria:

1. **Given** a customer has purchase history, **When** a new quotation is created, **Then** the AI suggests a product list based on their most frequently/recently purchased items with recommended quantities and last-sold prices.
2. **Given** the AI-generated draft, **When** the Sales user reviews it, **Then** they can: add new line items (search from catalogue), remove items, edit quantities, override unit prices, and add line-level or overall discount.
3. **Given** a product exists in the catalogue with a current price, **When** it appears in the quotation, **Then** both the catalogue price and the last-sold price to that customer are shown for reference.
4. **Given** no sales history exists for a customer, **When** a quotation is created, **Then** the system starts with a blank quotation and the user can add items manually from the catalogue with current catalogue prices.

FR-021: Quotation PDF Export

Field	Value
Priority	Must
Epic	Quotation Generation
Description	Generate professional quotation PDF with Evergreen branding. Includes: company header, customer details, quotation number (auto-generated), date, validity period, line items table (item, description, qty, unit price, total), subtotal, GST, grand total, terms & conditions, and authorised signatory. PDF can be downloaded or emailed directly to the customer.

Acceptance Criteria:

1. **Given** a completed quotation, **When** the user clicks "Generate PDF", **Then** a branded PDF is created with all quotation details, Evergreen logo, and professional formatting.
2. **Given** a generated PDF, **When** the user clicks "Email to Customer", **Then** the PDF is sent to the customer's email with a standard cover message.

3. **Given** multiple quotations for the same customer, **When** viewing quotation history, **Then** past quotations are listed with: quotation number, date, total amount, status (draft/sent/accepted/expired).

Notes: Quotation numbering format to be confirmed with Mrs Wee (e.g., QT-2026-0001).

FR-022: Quotation Management

Field	Value
Priority	Should
Epic	Quotation Generation
Description	List view of all quotations with status tracking. Filter by customer, date range, status (draft/sent/accepted/expired/rejected). Duplicate existing quotation to create new version. Track quotation-to-order conversion.

Acceptance Criteria:

1. **Given** a Sales user opens the Quotation list, **When** it loads, **Then** they see all quotations with columns: quotation number, customer, date, total amount, status, and last updated.
2. **Given** an existing quotation, **When** the user clicks "Duplicate", **Then** a new draft quotation is created with the same line items and pricing, ready for editing.
3. **Given** a quotation is accepted, **When** the user marks it as "Accepted", **Then** the status updates and it's linked to subsequent order/invoice tracking.

---

6. Non-Functional Requirements

Performance

Metric	Target
OCR processing time	< 30s per page
AI data extraction	< 15s per document
Content generation (text)	< 20s per product per platform
Image processing (bg removal)	< 10s per image
Page load time	< 2s
API response time	< 500ms p95
Concurrent users	10 (4 active + buffer)
Uptime	99.5%

Security

- ☑ HTTPS everywhere
- ☑ Email + password authentication with bcrypt hashing
- ☑ Role-based access control (Admin, Sales, Marketing)

- ☑ Secure file storage with signed URLs (no public bucket access)
- ☑ Session management with JWT (short-lived tokens + refresh)
- ☑ Input validation and sanitisation
- ☐ PDPA compliance review (Phase 2 — minimal PII in MVP)
- ☐ Audit logging for data access (Phase 2)

Accessibility

- Web app follows basic accessibility: semantic HTML, keyboard navigation, sufficient contrast
- WCAG 2.1 AA compliance is Phase 2

Scalability

- Expected: 4 users, ~50 documents/day, ~30 product listings/day
- Architecture supports 10x growth without re-architecture
- Stateless API server behind reverse proxy
- File storage: cloud object storage (S3-compatible)
- Database: PostgreSQL with connection pooling

7. Information Architecture / Sitemap

```
Evergreen AI Web App
├── Login
├── Dashboard
│   ├── Sales Overview (Sales/Admin)
│   └── Marketing Overview (Marketing/Admin)
├── Document Processing (Sales/Admin)
│   ├── Document Queue (list view)
│   ├── Document Review (side-by-side view)
│   └── Export History
├── E-commerce (Marketing/Admin)
│   ├── Content Generator
│   │   ├── Select Products
│   │   ├── Generate Content
│   │   └── Photo Enhancement
│   ├── Review Queue
│   │   ├── Pending Approval
│   │   ├── Approved
│   │   └── Published
│   └── Copy-to-Clipboard View
├── Product Catalogue (All roles)
│   ├── Browse / Search
│   ├── Product Detail
│   └── Import (Admin)
├── Settings (Admin)
│   ├── User Management
│   ├── Email Inbox Config
│   ├── WhatsApp Config
│   ├── Folder Watch Config
│   ├── ERP Export Format
│   └── System Logs
└── Help / About
```

8. Screen-by-Screen Specs

Screen: Document Queue

Field	Details
URL	/documents
Purpose	View and manage all ingested documents
Access	Sales, Admin

Key Elements:

- Filter bar: status dropdown, document type dropdown, source dropdown, date range picker
- Search bar: search by document number, company name
- Data table: columns — checkbox, doc type icon, doc number, date, company, total amount, source, status, confidence, received date
- Bulk actions bar (appears when rows selected): Export to Excel, Mark as Verified
- Status badges: colour-coded (red=failed, yellow=pending, green=verified, blue=exported)

Edge Cases:

- Empty state: "No documents yet. Documents from email, WhatsApp, and watched folders will appear here automatically."
- Error state: "Unable to load documents. Please try again."
- Large result set: server-side pagination, 50 per page

Screen: Document Review (Side-by-Side)

Field	Details
URL	/documents/:id/review
Purpose	Review and correct AI-extracted data against original document
Access	Sales, Admin

Key Elements:

- Left panel (50% width): document viewer — zoomable image/PDF, page navigation for multi-page docs
- Right panel (50% width): extracted data form
  - Header fields: doc type, doc number, date, company name, address (each in editable input)
  - Line items table: product name, matched SKU (dropdown with search), qty, unit price, line total — add/remove row buttons
  - Footer fields: subtotal, tax, grand total, payment terms, notes
  - Confidence indicators: yellow highlight on low-confidence fields
- Action bar: "Verify" (primary), "Skip" (secondary), "Flag Issue" (tertiary)
- Navigation: Previous/Next document buttons for sequential review

Edge Cases:

- No OCR text extracted: show original image + empty form with message "Automatic extraction failed. Please enter data manually."
- Very large document (20+ pages): lazy-load pages, show page thumbnails
- Unsupported format: show error with option to download original

---

Screen: Content Generator

Field	Details
URL	/marketing/generate
Purpose	Select products and generate marketplace content
Access	Marketing, Admin

Key Elements:

- Product selector: search/browse catalogue, select one or multiple products
- Platform selector: checkboxes for Shopee, Lazada (or both)
- Photo upload area: drag-and-drop or file picker for product photos (if not in catalogue)
- "Generate" button → triggers AI content generation
- Progress indicator during generation

Edge Cases:

- Product with no photo: prompt to upload one, or proceed with text-only generation
- Generation fails: show error, offer retry with different AI parameters

---

Screen: Content Review

Field	Details
URL	/marketing/review/:id
Purpose	Review and approve AI-generated content for a product
Access	Marketing, Admin

Key Elements:

- Product info header: name, SKU, category, source photo
- Photo section: original photo, background-removed version, enhanced versions, lifestyle shots — each with approve/reject toggle
- Platform tabs: Shopee | Lazada
  - Title: editable text input with character count
  - Description: rich text editor with formatted preview
  - Bullet points: editable list
  - Keywords/tags: editable tag input
- Action buttons: "Approve All", "Regenerate" (with optional instructions), "Reject"
- Copy buttons: "Copy Title", "Copy Description", "Copy All" — for clipboard paste to marketplace

Edge Cases:

- AI generates inappropriate content: reject + flag, regenerate

- Platform tab shows "not generated": offer to generate for that platform

---

9. API / Integration Requirements

Integration	Type	Purpose	Direction	Auth	Owner
Google Cloud Vision / AWS Textract	REST API	OCR processing	Outbound	API key	NexStack BE
Google Gemini / OpenAI GPT-4o	REST API	Data extraction + content generation	Outbound	API key	NexStack BE
GreenAPI	REST API + Webhooks	WhatsApp document ingestion	Bidirectional	API token	NexStack BE
Email (IMAP)	IMAP	Email document ingestion	Inbound	Email credentials	NexStack BE
Remove.bg / rembg	REST API / Library	Background removal	Outbound	API key	NexStack BE
OpenAI DALL-E / Stability AI	REST API	Lifestyle photo generation	Outbound	API key	NexStack BE
DigitalOcean Spaces	S3-compatible API	File/image storage	Outbound	Access key	NexStack BE
Email (SMTP / Resend)	SMTP / REST	Notification emails	Outbound	API key	NexStack BE

Key Integration Notes

- **ERP Integration:** None — Excel export only. The client imports the generated Excel file into their ERP manually.
- **Marketplace Integration:** Phase 1 is manual copy-paste. No Shopee/Lazada API integration in Phase 1.
- **LLM Usage:** Gemini as primary for data extraction (cost-effective for structured parsing). GPT-4o as fallback and for content generation (better creative writing).

10. Data Model Overview

```
User (1) ——— (N) ActivityLog
Document (1) ——— (N) DocumentLineItem
DocumentLineItem (N) ——— (0..1) Product
Product (N) ——— (1) Category
Product (1) ——— (N) ProductImage
Product (1) ——— (N) MarketplaceContent
MarketplaceContent (1) ——— (N) ContentImage
```

```
CatalogueImport (1) ———— (N) Product (via import batch)
SupplierProductMapping (N) ———— (1) Product
```

Key Entities:

Entity	Key Fields	Notes
User	id, email, name, role (admin/sales/marketing), password_hash, created_at	4 users, soft delete
Document	id, source (email/whatsapp/folder), source_ref, doc_type (invoice/po/do), doc_number, doc_date, company_name, subtotal, tax, total, status (queued/extracted/pending_review/verified/exported/failed), confidence_score, raw_text, original_file_url, created_at	Core entity for Workstream 1
DocumentLineItem	id, document_id, line_number, product_name, product_code_extracted, product_id (matched), quantity, unit_price, line_total, confidence_score	Line items from extracted documents
Product	id, sku, name, category_id, brand, description, unit_price, uom, image_url, status (active/inactive), created_at, updated_at	Shared catalogue
Category	id, name, parent_id	Hierarchical product categories
ProductImage	id, product_id, image_type (original/bg_removed/enhanced/lifestyle), image_url, created_at	AI-generated image variants
MarketplaceContent	id, product_id, platform (shopee/lazada), title, description, bullet_points (JSON), keywords (JSON), status (draft/approved/published), marketplace_url, created_by, approved_by, created_at	Generated listing content
ContentImage	id, marketplace_content_id, product_image_id, display_order	Links images to marketplace content
SupplierProductMapping	id, supplier_name, supplier_product_name, supplier_sku, product_id, created_at	Learned mappings for auto-matching
CatalogueImport	id, file_url, imported_by, products_added, products_updated, products_removed, imported_at	Import audit trail
ExportBatch	id, exported_by, document_ids (JSON), file_url, exported_at	Excel export audit trail

ActivityLog	id, user_id, action, entity_type, entity_id, details (JSON), created_at	Audit trail

11. Error Handling & Edge Cases

Scenario	User Message	System Action	Severity
OCR fails on document	"Extraction failed for this document. You can enter data manually."	Mark as "failed", show in queue with warning icon	High
LLM service unavailable	"AI processing is temporarily unavailable. Your documents are queued."	Queue for retry every 60s, process when restored	High
Email inbox connection lost	(Admin notification) "Email monitoring disconnected. Check credentials."	Retry every 5 min, alert admin after 3 failures	High
WhatsApp webhook down	(Silent)	Queue messages, retry, alert ops	Critical
Unsupported file format	"This file format is not supported. Please use PDF, JPEG, or PNG."	Move to unsupported folder, log	Low
Catalogue import validation error	"Row 15: Missing product name. Row 23: Invalid price format."	Show per-row errors, allow fix and re-upload	Medium
Content generation fails	"Content generation failed for this product. Click Retry."	Log error, allow retry with different prompt/model	Medium
Background removal fails	"Could not process this image. Try a different photo."	Fallback to original image, allow manual upload	Medium
Excel export format mismatch	"Export format error. Contact admin to update ERP template."	Log error, notify admin	Medium
Storage quota exceeded	(Admin alert) "Storage is running low."	Alert admin, suggest cleanup/archival	High

12. Dependencies & Risks

Risk	Likelihood	Impact	Mitigation	Owner
Poor scan quality reduces OCR accuracy	High	Medium	Allow manual entry fallback, document quality guidelines for client, image pre-processing (deskew, contrast)	NexStack BE
ERP export format undocumented / changes	Medium	High	Get sample files early, build configurable template, test with actual ERP import	NexStack BE + Client

Diverse document formats from suppliers	High	Medium	LLM few-shot examples per supplier, learning from corrections, supplier-specific templates over time	NexStack BE
AI-generated marketing content quality inconsistent	Medium	Medium	Human review queue ensures quality gate, iterative prompt tuning, allow regeneration	NexStack BE
Lifestyle photo generation produces poor results	High	Medium	Mark as experimental, focus on bg removal + enhancement as core, lifestyle as bonus	NexStack BE
Product catalogue incomplete/outdated	Medium	High	Regular sync reminders, diff view on import, handle unmatched products gracefully	Client + NexStack
WhatsApp API rate limits or policy changes	Low	High	GreenAPI abstraction, message queue, monitor policies	NexStack BE
Budget constraints limit AI API costs	Medium	Medium	Use Gemini (cheaper) for extraction, cache results, batch processing, monitor API spend	NexStack BE

13. Release Criteria / Definition of Done

- ☐ All Must-have features (FR-001 through FR-015) implemented and tested
- ☐ OCR + AI extraction tested with 20+ real documents from client (invoices, POs, DOs)
- ☐ Side-by-side review UI functional with edit + verify flow
- ☐ Excel export tested with actual ERP import (client confirms format works)
- ☐ Product catalogue import/sync working with client's real data
- ☐ Content generation tested for 10+ products on both Shopee and Lazada formats
- ☐ Background removal working on product photos
- ☐ All 4 user accounts created with correct role-based access
- ☐ Email and WhatsApp ingestion tested with real sources
- ☐ End-to-end flow tested: document arrives → OCR → review → export
- ☐ End-to-end flow tested: select product → generate content → review → approve → copy
- ☐ Performance benchmarks met (page load < 2s, OCR < 30s)
- ☐ Basic error handling for all scenarios in error matrix
- ☐ UAT sign-off from client
- ☐ Deployed to production environment

14. Tech Stack Recommendations

Layer	Technology	Rationale

Frontend	Next.js 14 (App Router) + Tailwind CSS + shadcn/ui	Modern React framework, great DX, good for internal tools
Backend	Next.js API Routes or Hono.js	Keep it simple — 4 users, low traffic
Database	PostgreSQL (Supabase or DigitalOcean Managed)	Reliable, good JSON support for flexible fields
OCR	Google Cloud Vision API	Best accuracy for mixed English/Chinese documents
AI/LLM	Google Gemini (extraction) + OpenAI GPT-4o (content)	Cost-effective extraction + high-quality content
Image Processing	remove.bg API + Sharp.js	Background removal + image manipulation
AI Images	OpenAI DALL-E 3	Lifestyle shot generation
File Storage	DigitalOcean Spaces (S3-compatible)	Cost-effective, Singapore region available
WhatsApp	GreenAPI	Proven in PilotNow project
Email Ingestion	IMAP client (imapflow)	Simple, reliable email monitoring
Hosting	DigitalOcean App Platform or VPS	Singapore region, cost-effective
Auth	NextAuth.js or custom JWT	Simple for 4 users

15. Timeline & Milestones

Week 1: Foundation & Document Pipeline

- Project setup, CI/CD, database schema
- User auth + role-based access
- Document ingestion: email, WhatsApp, folder monitoring
- OCR integration (Google Cloud Vision)
- Basic document queue UI

Week 2: Document Review & Catalogue

- AI data extraction (LLM structured parsing)
- Side-by-side review UI (original + extracted data)
- Product catalogue import (CSV/Excel)
- Catalogue search/browse UI
- Product matching logic (exact + fuzzy)

Week 3: E-commerce Content & Export

- Content generation: titles, descriptions, bullet points (Shopee + Lazada)
- Photo enhancement: background removal, image optimization
- Content review queue UI
- Excel export in ERP format
- Copy-to-clipboard functionality

Week 4: Polish, Testing & Launch

- Dashboard with key metrics
- Activity logging
- End-to-end testing with real client data
- Bug fixes and UI polish
- UAT with client
- Production deployment
- Handover documentation

Budget: SGD 50,000 **Timeline:** 4 weeks **Team:** NexStack development team

16. Appendix

Client Details

Field	Value
Company	Evergreen Group Pte Ltd
Contact Person	Mrs Wee
Phone	+65 9683 1343 / +65 9424 4633
Email	mrswee@evergreen.com.sg
Address	8 New Industrial Road, #01-02/03, LHK3, Singapore 536200
Website	https://evergreen.com.sg
E-shop	https://eshop.evergreen.com.sg

- Marketplace presence: Shopee.sg, Lazada.sg
- Reference project: PilotNow (WhatsApp integration, GreenAPI, similar tech stack)
- Total users: 4 (2 sales, 2 marketing)
- ERP: Custom system with Excel import only (no API)