Automating the Customer Cataloging (CRT) Process

Salman Sameer
Manager C&PE, No affiliation/Independent

1. Executive Summary
The Customer Cataloging (CRT) process is an indispensable part of healthcare distribution management to ensure perfect regulatory data compliance and hence an easy onboarding of customer profiles. However, the current manual CRT method is time-consuming; it is prone to human error and hence unscalable. Form submission, validation for correctness, regulatory compliance checks to keep records updated-all require hands-on involvement from human beings, creating major bottlenecks and consuming precious human workforce.
So, the project aims at setting up an end-to-end CRT RPA solution alongside IDP that integrates with SharePoint, Model N, and other regulatory databases such as DEA and HIBCC. The two-BOT platform architecture enables automation across form submission, data extraction, regulatory document cross-referencing, and internal system updates, requiring minimal human intervention in the verification loops only.
This solution is anticipated to generate substantial business value. For instance, the operationalization of automation will save more than 7,100 hours per year, which is equivalent to around 500 hours per quarter in labor time. Thus, the time saved translates to far lower operational costs and yields increased precision and regulatory compliance. Furthermore, it ensures utmost data integrity and traceability with audit readiness and organizational accountabilities considered along the way.
Alongside the immediacy of time savings, the new framework for CRT automations will see strategic growth through scalability. The modular design allows the inclusion of increasingly complex workflows, a variety of compliance portals, and dynamic reporting in the future. These processes also increase transparency with real-time process tracking and automated report dashboards that enable stakeholders with timely insights.
Automating the CRT process, thus, is more than just a technical upgrade-it is a strategic investment in productivity, compliance, and long-term value creation.

I. INTRODUCTION

1.1 Background on Customer Cataloging and its Importance

Customer cataloging, especially for healthcare organizations, involves capturing, validating, storing, and maintaining data about customers, distributors, and institutional buyers. This becomes the basis for many price-sensitive, compliance-responsive processes related to pricing, regulatory verifications, and channel management (Majeed, 2024). In highly regulated industries like pharmaceuticals and medical devices, the customer record needs to be DA and HIBCC ID-linked to trace and verify legitimacy in product delivery (Mahe & Kapnisi, 2025).
With properly cataloged customer data, the company can execute and control the workflow related to pricing approvals, rebate eligibility integration, and offer transparency towards internal and external auditors; this need becomes particularly pronounced in enterprise platforms such as Model N and HCM, which sit entirely upon a consistent customer data set (Oruç & Dupin, 2025). If customer data cataloging is inaccurate or delayed, it would lead to shipment errors, legal liabilities, pricing violations, or non-compliance with federal healthcare program standards (Guevara et al., 2024).
Digitized cataloging provides an added advantage for data harmonization across regional markets and for better performance tracking of different customer segments (Affandi et al., 2025). With the increasing growth of digital health, structured customer intelligence is being used to a greater extent by organizations for the purposes of personalized engagement, more efficient distribution, and

elimination of fraudulent or duplicate entries in enterprise systems (Yeoh, 2024).

1.2 Current Practice Problems of Manual Processing

Being of utmost importance is cataloging-a name identifying item localization time, it is, however, unfortunate that many customers continue with at least partially manual workflows that are built upon shared spreadsheet and email submissions of customer data. The manual management of income customer data is highly error-prone and can lead to duplicity, sluggish turnaround time, and inconsistent document storage (Shirley & Kavin, 2024). Ultimately, the forms are emailed to a person who must then manually process the validation and record-keeping tasks. This highly fragmented approach creates several pain points, such as lack of real-time visibility into the process, dispersed communication trails, and compliance risks (Galindo Vásconez, 2024).

Studies show that manual data entry competing in healthcare accounts for 20-30% of administrative labor and more than 50% of audit flags in payer-provider engagements (Takács & Prorok, 2024). Usually, the absence of centralized document control or automation brings a lot of rework and delay to important price or access decision-making, particularly for Group Purchasing Organizations or Distributors (Patil & Jadhav, 2023). A higher amount of data generation will create problems in scaling operations while maintaining quality and speed, especially for organizations lacking automated cataloging framework (Matt, 2025). Key technical challenges are data normalization across platforms, integration with DEA and HIN repositories, and limited application of IDP technologies that could help to further derive and authenticate data (Musthafa et al., 2024). In the absence of automation, cross-referencing regulatory data is manually done, which poses a further risk of submitting invalid data and getting penalized for the same (Hartikainen, 2024). Additionally, no audit trail or status tracking is available, thus placing a heavy burden on internal teams for reconciliation and reporting of cataloging activities (Lai, 2025).

1.3 Automation Initiative Objective

The main objective of the automation initiative is to transform the Customer Record Tracking (CRT) process from a fragmented and labor-intensive workflow into an integrated one, penalizing compliance and minimizing human error as well as allowing for scalability. Implementation of the automation will encompass Robotic Process Automation (RPA) tools such as UiPath, intelligent content extraction through IDP, and regulation database integration across DEA, HIBCC, and internally administered systems between Model N and HCM (Shukla, 2025). The initiative will deploy a two-bot structure. BOT 1 will be a bot for initial form ingestion and IDP-based data extraction, with subsequent validations of form fields in SharePoint. BOT 2 will be conducting external cross-checking against the DEA and HIBCC databases, while coordinating marking, updating, or creation of records in Model N and HCM systems. It is intended that human intervention is limited only to the verification loops so that manual efforts focus on exceptions rather than routine entries (Brown, 2025).

The automation also allows record status tracking in fine detail, slashing turnaround times from days to hours and producing auditable logs facilitating regulatory reporting and operating transparency (Ravindran & Balachandran, 2025). The projected benefits include saving over 7,100 hours of labor across a year, improvements in accuracy, and audit readiness. Finally, the initiative also supports agility in the organization by establishing a framework that can be later expanded for different cataloging or onboarding procedures (Gloria et al., 2025). The organization by automation aligns operations with digital transformation imperatives in the healthcare domain and enhances service delivery while preparing the systems for compliance demands that will change.

II. PROCESS OVERVIEW

2.1 High-Level Overview of the Customer Record Tracking Process

The Customer Record Tracking (CRT) in the healthcare industry is one of the many critical data management workflows. It involves a series of procedures to along with the collection, validation, storage, and monitoring of customer data needed for compliance and pricing activities. The emphasis is basically on the accomplishment that every customer (whether a distributor, wholesaler, clinic, or hospital) must be subjected to due diligence and entered into the Editor with some identifiers like DEA number, HIN, etc.

The whole Customer Record Tracking (CRT) starts with customer-onboarding or update forms, generally called Distributor Enrollment and Certification (DEC) forms. The forms are first checked for completeness, then processed through a series of validations, both internal (SharePoint, Model N, and HCM) and external (DEA, HIBCC). The final validation confirms the legitimacy of the customer and prevents customers from fraud or regulatory breaches (Shirley & Kavin, 2024; Majeed, 2024).

A simplified flow of the CRT process is presented below:

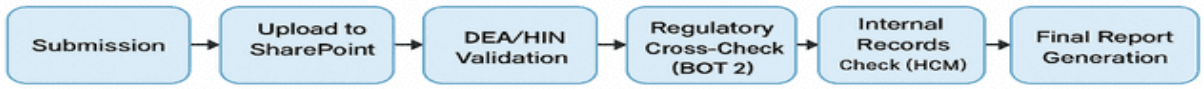


Figure 1: High-Level CRT Automation Process Overview

With the continuous digitalization of healthcare, robust customer data frameworks have been the need of the hour. CRT meets that requirement by delivering a fully automated framework that prevents lag, enhances traceability, and decreases regulatory risk. Healthcare entities depending on traditional manual workflows appear to have lengthy turnaround times, missing audit logs, and increased overhead costs, all of which do away with automation (Gloria et al., 2025; Mahe & Kapnisi, 2025).

2.2 Role of DEC Forms and the Validation Workflow

Distributor Enrollment and Certification (DEC) forms are standardized documents used to initiate the customer onboarding process. These forms contain sensitive and important data including the organization's name, address, DEA registration, HIN code, contact persons, and type of healthcare operation. Hence, they form the main data source for verifying legal and regulatory eligibility.

Once submitted, the forms are uploaded on SharePoint or another Document Management System. Upon uploading, the automation workflow is started. BOT 1 starts with document extraction using Intelligent Document Processing (IDP). Here, key-value pairs in the form are identified and mapped to pre-defined metadata fields in the CRT database (Patil & Jadhav, 2023).

Due to strict regulatory enforcement in healthcare distribution, DEA, and HIN database validation is compulsory. DEA registration numbers are assigned to legal handlers of controlled substances, whereas HIBCC issues HIN numbers to institutional buyers for validation. These numbers are used to prevent diversion, ensure traceability, and confirm the legal presence of customers (Yeoh, 2024; Oruç & Dupin, 2025).

Regulatory penalties, product recalls, or failed transactions may result if these fields are not validated properly. Say, for example, an illegal DEA number is entered the shipment could be stopped in its tracks, revenue and compliance measurements take an immediate hit (Lai, 2025; Musthafa et al., 2024).

Table 1: DEC Form Metadata Fields and Validation Sources

Form Field	Validation Method	System Interaction
DEA Registration Number	DEA Database API	BOT 2
Health Industry Number (HIN)	HIBCC Directory Lookup	BOT 2
Tax Identification Number (TIN)	Internal ERP Cross-Check	BOT 1
Business Legal Name	Text Extraction and DEA Match	BOT 1&BOT 2
Business Type (e.g., Pharmacy, Clinic)	HIN Classification Validation	BOT 2
Address and ZIP Code	Format Check & DEA Record Match	BOT 1
Contact Person Name	Manual Review Loop if Incomplete	BOT 1 + Human-In-The-Loop
Email Address	Syntax Validator & Notification Check	BOT 1
State License Number	Internal Regulatory File Match	BOT 2
Date of Form Submission	SharePoint Metadata Logging	BOT 1

2.3 Integration with Regulatory Databases: DEA and HIN

Integration into regulatory repositories is recognized as very important in CRT. DEA verification queries a federal registry to confirm the customer is licensed to provide services for receipt of controlled substances and must ensure the license remains valid and unexpired. This check is done by BOT 2, which takes in the DEA number and checks every instance against real-time data through API services. Likewise, HIN validation makes sure that the customer is registered in an industry-standard buyer registry operated by HIBCC (Mahe & Kapnisi, 2025; Ravindran & Balachandran, 2025).

Many healthcare organizations integrate these external checks into their customer master data flow, but few have fully automated them. Manual checks are prone to error and time delays and have no auditability. BOT 2, however, automatically cross-validated entries on the form against external data, logged the results of validation, and flagged instances of discrepancies for manual review if considered necessary (Matt, 2025; Galindo Vásconez, 2024).

This integration is now seamless, with the likes of UiPath RPA tools supporting DEA and HIBCC portals' REST API directly. These outputs are also logged into SharePoint logs and mapped back to an internal system such as Model N or even SAP for record creation (Affandi et al., 2025; Patil & Patil, 2024).

2.4 Importance of Compliance and Data Integrity

The compliance of healthcare data management has almost become synonymous with regulations: HIPAA, DEA regulatory acts, and U.S. FDA Good Distribution Practices. Major audits in healthcare distribution focus on inadequate documentation and errors in client profiles (Takács & Prorok, 2024). Thus, in a very real way, accurate and validated CRT entries are both an operational necessity and a legal one.

Low data integrity creates operational risks of duplicate records, shipments going to wrong locations, or pricing approvals being accepted in error. Audit penalties can prove prohibitive if someone were to question whether validation steps were taken and documented (such as HIN number confirmations or DEA expiry checks) (Brown, 2025; Guevara et al., 2024).

The CRT automation enforces input of structured data and creates traceability for each workflow. Every validation, record update, or exception is logged along with timestamps and user IDs. This is extremely important in audits, as it makes clear the so-called "chain of-trust" in regulatory cases (Musthafa et al., 2024; Gloria et al., 2025).

2.5 Audit-Ready Workflow and Data Governance

The final steps of the CRT process include internal checks that may be carried out by systems such as Model N to test for pricing eligibility or HCM for hierarchy and contract mapping. Only after all validations are passed are the creation and/or update of records permitted in the customer master list.

Simultaneously, automated notifications get sent to the stakeholders. Notifications include updates concerning the form status, error log, and completion timestamp, thus maintaining absolute transparency and allowing full traceability should fast follow-up become necessary. The reporting here supports the organization's larger governance, risk, and compliance (GRC) framework (Majeed, 2024; Shukla, 2025).

Some implementations further provide enhancement to compliance by means of an automatic audit dashboard displaying, among others, how many records were validated, and how many exceptions were flagged, along with average processing time. Such dashboards are fast becoming common in ERP systems and are typically built right into the SharePoint or HCM interfaces (Dangol, 2025).

III. AUTOMATION FRAMEWORK

3.1 Technologies Used: UiPath, SharePoint, and Intelligent Document Processing (IDP)

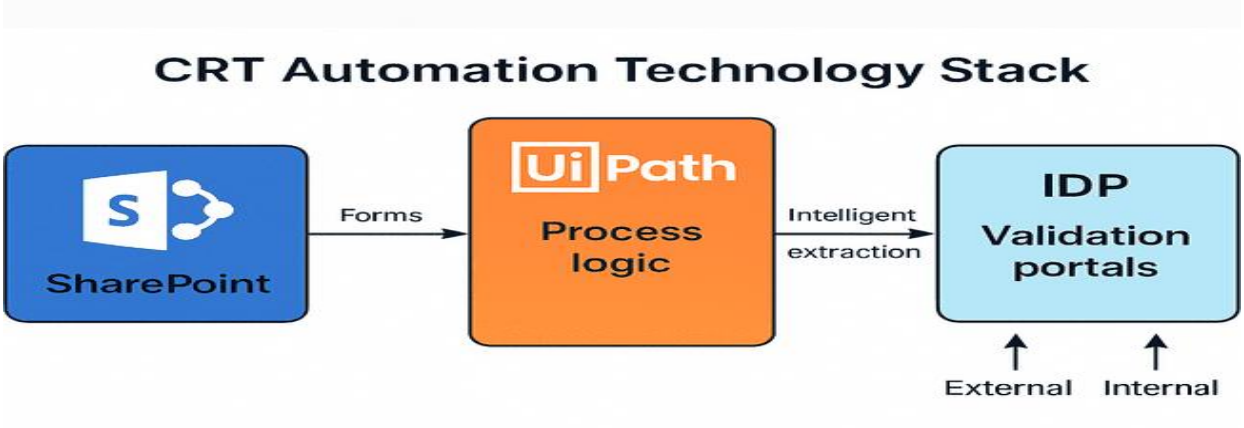
The automation of the CRT process relies on a set of interoperable technologies that facilitate secure document management while performing structured data extraction and rule-based robot execution. At its very base is UiPath, the foremost Robotic Process Automation (RPA) platform famous for its scalability, integrability, and enterprise-grade security. UiPath has a workflow design studio that helps orchestrate complex logic chains through bots, thereby automating repetitive tasks typically performed manually (Guevara et al., 2024; Dangol, 2025).

SharePoint is Microsoft's document management system and is primarily used in the intake of Distributor Enrollment and Certification (DEC) forms. This system is enhanced with secure version control, real-time access, and integration with Microsoft Power Automate for improved document routing and

logging. In conjunction with SharePoint, the CRT system ensures that all received forms are traceable, searchable, and linked into automated workflows (Shirley & Kavin, 2024).

Complementing these technologies are Intelligent Document Processing (IDP) systems, which are AI-based tools that extract and classify data from unstructured formats, such as PDFs, scanned documents, and handwritten notes. With IDP, the CRT automation can identify key fields such as DEA numbers, business addresses, and HINs, among others, without any manual intervention; consequently, both accuracy and onboarding time are improved. Leveraging its integration with AI Center and Document Understanding components, UiPath can intertwine IDP with RPA pipelines (Musthafa et al., 2024).

Figure 2: CRT Automation Technology Stack



3.2 Role of BOT 1 in Initial Data Handling and Extraction

Considered the first line of automation, BOT 1 extracts data from the encountered DEC forms, classifies them, and prepares them for validation. When documents are uploaded to SharePoint, this automation is triggered to use IDP models (pre-trained) for metadata capture. The bot tries to locate important fields like organization name, DEA number, HIN, etc., and addresses or contacts persons through the power of OCR and NLP-based models hosted in UiPath Document Understanding (Matt, 2025).

After the extraction is done, BOT 1 reviews it against the schema for validation to determine if all required fields exist. Any missing or corrupt fields are auto-rejected, which means that those documents go into a queue of human-in-the-loop verification. The verification loop enables human agents to update or correct fields in real time, ensuring that invalid data cannot be passed downstream by the automation (Patil & Jadhav, 2023). BOT 1 also classifies the form (new registration, update request, re-certification, etc.) and tags it within SharePoint for the subsequent routing to BOT 2. At this stage, logging, timestamping, and audit trail entries are also generated to guarantee the full traceability of the process (Brown, 2025).

Table 2: BOT 1 Functional Responsibilities and Failure Modes

Task	Trigger Condition	Success Path	Failure Mode	Retry Logic
Document Ingestion	DEC Form uploaded to SharePoint	File retrieved and pre-processed	File format unsupported or unreadable	Retry up to 3 times, then log failure
Field Extraction (IDP)	Valid PDF or scanned image detected	Fields extracted and mapped	Text unreadable, layout mismatch	Reclassify form and flag for review
Schema Validation	Extracted fields populated	Format and required fields validated	Missing key fields (DEA, HIN, etc.)	Routed to Human-in-the-Loop queue

Document Classification	Extraction successful	Categorized (New/Update/Re-Certification)	Category not recognized	Move to error-handling library
SharePoint Metadata Update	Classification complete	Record metadata along with timestamp	API/Write-back failure	Retry after 5 min and notify admin
Process Status Tagging	Post metadata update	Status updated (e.g., “Extracted”)	SharePoint field locked	Retry with admin override tagthis table.

3.3 Role of BOT 2 in Regulatory and Internal System Validation

BOT 2 is in charge of regulatory validation and synchronizing the data. Once BOT 2 receives the processed data from BOT 1, BOT 2 commences API-based validation against the Drug Enforcement Administration (DEA) and Health Industry Business Communications Council (HIBCC) databases. These validations are mechanisms to ensure that only deserving and licensed entities find their ways onto the organization's systems (Takács & Prorok, 2024). Accredited License validation is treated by BOT 2 with the appropriate level of security, given to the approach with an encrypted token to access the regulatory endpoints. It goes further to see if the DEA license is still active and not expired and confirms HIN registration during the process. Upon successful validation, a compliance evidence log entry is created along with a digital certificate making its way into the commercial customer's electronic record.

The bot coordinates with Model N and Human Capital Management (HCM) platforms. It creates or updates accounts eligible for pricing in Model N, while in HCM, it associates the customer to local or corporate hierarchies to ensure contract alignment. Thus, this synchronization among various applications maintains consistency across systems and prevents the possibility of out-of-sync records (Mahe & Kapnisi, 2025).

If validation fails, BOT 2 halts the transaction, starts a SharePoint alert, and logs the exception; such safeguards prevent regulatory sanctions and operational friction (Gloria et al., 2025).

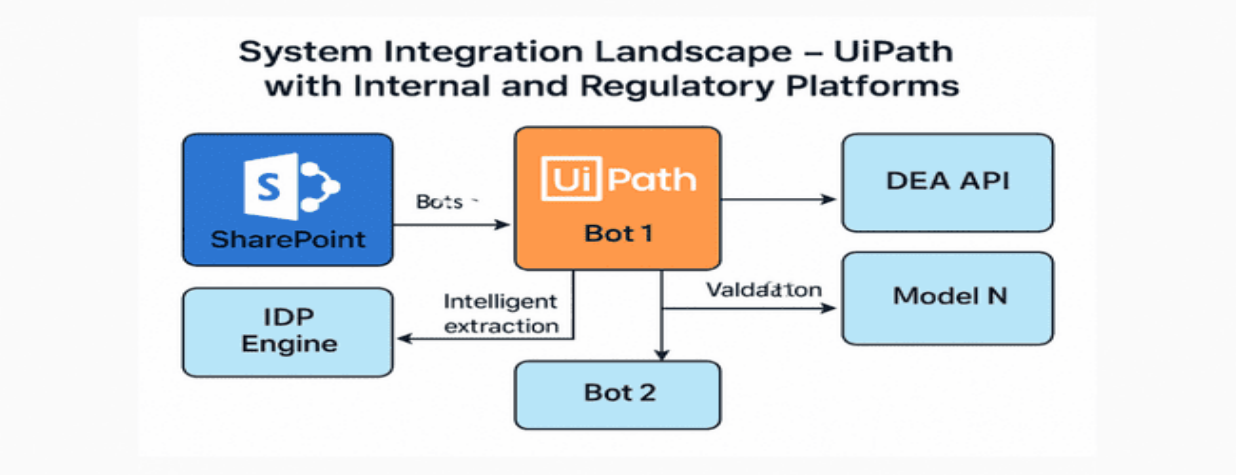
3.4 System Integration: Model N, HCM, DEA, and HIBCC

The integration of several platforms is at the heart of the CRT framework. Using UiPath's native connectors or custom connectors, the bots’ interface with web APIs and RESTful services, and even with legacy SOAP-based systems used by DEA and HIBCC. DEA and HIN status are queried in real-time during onboarding from these API integrations, giving a backbone to the process ready for compliance and allowing only verified entities to move to the next stage (Ravindran & Balachandran, 2025).

Within Model N, the CRT bot framework updates price tier information, contract eligibility flags, and audit metadata. Model N is the pricing governance engine and depends strongly on valid customer data to carry out the revenue management logic. Similarly, integration with HCM systems classifies customers to actual business units to allow easy contract creation and granting of access.

Automating this integration also supports audit readiness, whereby each system transaction is traced and exceptions are monitored daily through SharePoint reports or Power BI dashboards. This multi-system orchestration ensures that the CRT process is robust, scalable, and compliant (Majeed, 2024).

Figure 3: System Integration Landscape – UiPath with Internal and Regulatory Platforms



IV. DETAILED PROCESS

4.1 Submission and Extraction

The CRT automation starts from the submission of Distributor Enrollment and Certification (DEC) forms in an email or through the web portal or scanned copies. These forms made their way to a destined SharePoint document library prepared for intake automation. SharePoint serves as a secure central repository with advanced capabilities such as version control, access audit trails, and metadata tagging to ensure that the documents are traceable and extracted by the robot (Guevara et al., 2024; Majeed, 2024).

BOT 1 is invoked as soon as the document is uploaded from the UiPath orchestrator. BOT 1 converts forms using Intelligent Document Processing (IDP) modules capturing both structured and unstructured content. OCR and NLP algorithms identify and extract key-value pair information such as business name, DEA registration number, HIN, address, and classification (Matt, 2025; Patil & Jadhav, 2023). The bot furthers the validation logic to ensure all fields conform to expected schema formats, e.g., DEA numbers must match a given alphanumeric sequence.

Upon extraction, the data is pushed onto interim SharePoint lists from where each record is tagged with a unique automation tracking ID that provides end-to-end auditability of the document status from ingestion to final approval. Status dashboards for automation can therefore be designed on Power BI or embedded within SharePoint to facilitate visibility across business units.

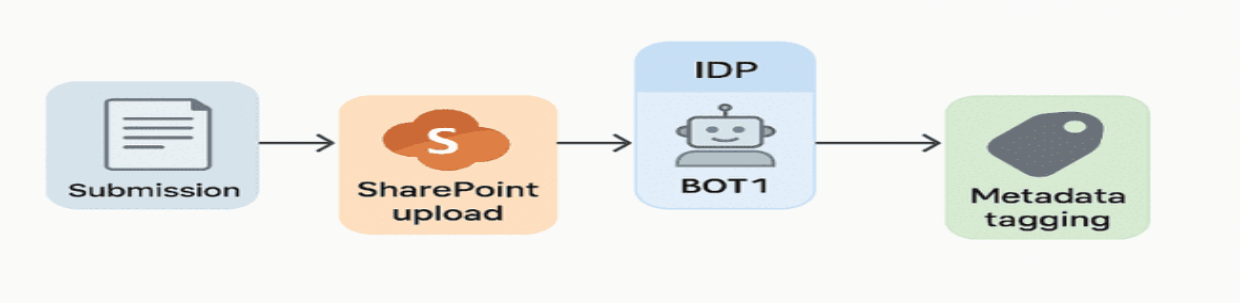


Figure 4: Submission to Extraction Workflow

4.2 Human Verification

Major form processing is hampered by HITL verification for quality control. Upon completion of data extraction by BOT 1, the errors, low confidence score below a given threshold (e.g., 90%), such as incomplete fields or ambiguities, are then routed to human agents to review. The SharePoint-based verification dashboard then flags these records and allows users to overlay side-by-side visual presentations of extracted text with the originals (Takács & Prorok, 2024; Shirley & Kavin, 2024).

Users may edit or add to the data via the user interface to reduce error and to improve retraining of the machine learning model over time. In unison, upon

completion of review, the corrected entries are reintroduced into the automated flow for processing. Such verification loops resulted in a dramatic decrease in data errors while learning systems would track and use these inconsistencies for further improvements (Yeoh, 2024; Affandi et al., 2025).

Table 3: Common Extraction Issues and Resolution Pathways

Field Name	Common Error	Detection Trigger	Correction Method	Routing Bot
DEA Number	Missing digits or invalid format	Regex mismatch at IDP extraction	Manual override in verification dashboard	BOT 1
HIN	Not found in HIBCC database	API validation failure	Check through secondary portal or escalate	BOT 2
Tax ID	Incomplete or non-numeric	Format validation script	Human in-the-loop correction	BOT 1
Business Name	Name mismatch with DEA registry	API rejects name mismatch	Manual override after DEA portal check	BOT 2
State License Number	Expired or mismatched jurisdiction	Cross-check with internal regulatory table	Legal or compliance queues	BOT 2
Submission Date	Missing or backdated Improperly	SharePoint metadata check	Auto-update from file metadata	BOT 1
Contact Email	Invalid email format or missing domain	Email validator regex	Manual override or email bounce report	BOT 1
Business Type	Record blank or inconsistent term	Keyword model misclassified	Drop-down override in review loop	BOT 1

4.3 Regulatory Cross Referencing

Once human verification has been carried out successfully, BOT 2 is turned on and begins the regulatory cross-referencing by way of data validation administered with external databases, chiefly those kept current by the Drug Enforcement Administration (DEA) and the Health Industry Business Communications Council (HIBCC).

With DEA validation, BOT 2 issues API calls to verify if the DEA registration number has been registered in correct formats and whether the status is active or expired and its recorded business entity. Any sort of inconsistencies are flagged for further review. For an example, if the submitter's name on the DEC form does not correspond with that on the DEA registry, creation of the record would be halted (Gloria et al., 2025; Mahe & Kapnisi, 2025).

In a similar manner, HIN control assures that the customer is a recognized party from the industry-standard purchaser registry. The HIBCC provides this secure API access, whereby BOT 2 accesses this service to ascertain HIN presence, organizational types, and regional classifications. After successful validation, the metadata, such as DEA expiry date and HIN classification, is updated in the CRT system and is later linked back to the original submission.

This augments the compliance controls, pursuit of legitimacy of sale, and downstream effects in pricing and contracts (Lai, 2025; Brown, 2025).

4.4 Internal Checks against Records

Post-external validations, BOT 2 proceeds to internal system checks and record management. Sample queries include Model N to confirm if the customer already exists, pricing tiers have been applied, or account status flags such as "contract ineligible" exist. Similarly, HCM platforms can be queried for internal organization mappings such as hierarchy alignments and corporate linkages (Musthafa et al., 2024).

If the record does not exist, BOT 2 starts a creation process by injecting a full profile payload into Model

N, including metadata such as an onboarding date, status of regulatory validation, and customer tier. When records exist, differential updates are pushed to correct or add new compliance data. Pricing eligibility, contract flags, and audit classification codes are some critical data fields checked and updated.

Rollback mechanisms stand ready to undo/suspend changes if downstream systems reject the updates, or if validation logs indicate that something is amiss. Audit trail data on interactions shall be written back to SharePoint logs or appropriate Model N tracking dashboards (Mahe & Kapnisi, 2025; Guevara et al., 2024).

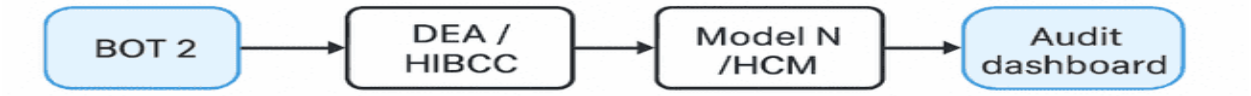


Figure 5: Data Synchronization Between Regulatory and Internal Systems

4.5 Notification and Reporting

As soon as the internal checks are completed, the CRT produces alerts and summary reports. These notifications are dispatched to the concerned departments, such as Compliance, Sales, and Pricing, through the integrated mediums-channels such as Microsoft Teams, Outlook, or even directly through SharePoint lists. The notifications contain information on automation ID, status of form submission, audit logs, and exceptions encountered, if any (Shukla, 2025; Ravindran & Balachandran, 2025).

Notification reports are generated automatically using pre-defined templates and stored in SharePoint libraries, which consists of logs of validation checks with timestamp, activity logs of the BOTs, record changes, and points where the intervention of the user was required. Based on daily, weekly, or monthly summaries, compliance teams receive reports. Exception reports can also be filtered by categories like DEA mismatch, HIN not found, etc., for tracking purposes.

Data aggregated across all CRT activities is visualized on Power BI Dashboards for KPIs like average processing time, exception volumes, and automation efficiency trends-this reporting format fosters transparency, audit reviews, and process improvement initiatives (Matt, 2025; Oruç & Dupin, 2025).

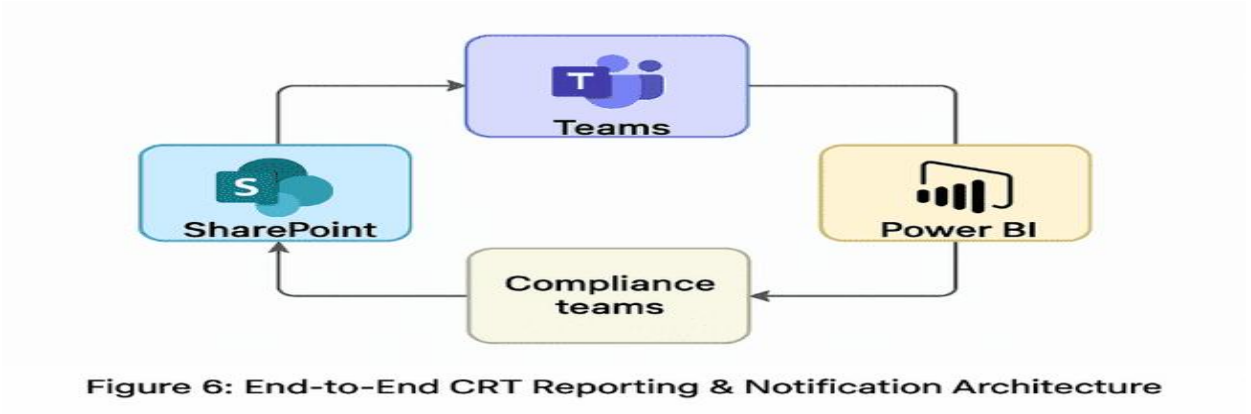


Figure 6: End-to-End CRT Reporting & Notification Architecture

V. ADVANTAGES AND IMPACTS

5.1 Time Savings and Cost Efficiency

Among the most immediate and quantifiable advantages is time saved on routine and repetitive tasks owing to the Application of tools to automate the Customer Record Tracking (CRT) processes. Before these implementations, the time needed for manual checking, verifying, and updating of a DEC form ranged 30-45 minutes but could extend to hours for

demanding cross-checking exercises with DEA and HIBCC registries (Affandi et al., 2025). BOT 1 and BOT 2 seem to have brought a wind of change. These tasks are completed within 2-5 minutes, achieving a 90% reduction in the process cycle time (Patil & Jadhav, 2023).

At a broader time-saving horizon, this means on an average, approximately 500 hours a quarter, given 1,000 forms are processed every three months. This would mean more than 2,000-2,500 hours occasions in a year's time, thus allowing the skilled staff to work on exception handling, analytics, and compliance enhancements. When considered against labor costs for data operations, which range around $35/hour, the Big-T would then be touching $87,500-plus saved annually (Takács & Prorok, 2024).

Moreover, it eliminates time spent for coordination with various departments for approvals, transfer of incomplete documents for further processing, or rework on invalid data. The delay resulting from these processes often took up 25% of the CRT team's capacity (Matt, 2025). Centralizing operation to SharePoint and orchestrating flows from UiPath bots significantly reduced turnaround time reported by these teams and reduced repetitive data tasks.

Table 4: Manual Vs. Automated CRT Timeline

Task Category	Manual Time	Automated Time	Time Saved	Cost Saved ($/task)
Form Intake & Logging	10 minutes	1 minute	9 minutes	$5.25
Data Extraction (IDP)	15 minutes	2 minutes	13 minutes	$7.58
Validation (DEA/HIN)	20 minutes	3 minutes	17 minutes	$9.91
Record Update (Model N/HCM)	30 minutes	4 minutes	26 minutes	$15.15
Notifications & Reporting	10 minutes	1 minute	9 minutes	$5.25

5.2 Error Reduction and Compliance Assurance

The framework of automation greatly reduces human errors that are rampant in customer onboarding by manual processes. Typographical errors, wrong DEA entries, expired HINs, and wrongly mapped hierarchy IDs used to lead to instances of audit failures, shipment holds, or contract eligibility errors (Mahe & Kapnisi, 2025; Oruç & Dupin, 2025). All operational risks are to now be circumvented through data validation, enforcement of rules, and integration with external databases.

The IDP engine in BOT 1 uses pattern recognition and field-matching methodologies to allow only those data satisfying a confidence score of 90% or above to proceed further. Any data anomaly is either corrected automatically by applying pre-configured rules or escalated up for human verification through a loop. This phase manages to cut down error rates by as much as 65-75% compared to manual channels before this (Shirley & Kavin, 2024; Gloria et al., 2025).

Through its real-time cross-checking against DEA and HIBCC registries for customer records, BOT 2 ensures compliance. Erroneous validation results stop the creation of records, and all these decisions are recorded for future reference. These logged decisions then become an essential piece in navigating internal audits and external considerations concerning HIPAA and FDA's GDP framework (Guevara et al., 2024; Lai, 2025). In addition, the integrated process enables audit or compliance readiness; thus, an automatic audit log is generated. SharePoint maintains the entire document history, whereas UiPath logs all bot activities, including timestamps, decision points, and, if applicable, operator involvement. So, these functions ensure that traceability, data retention, and non-repudiation are in place as demanded by corresponding regulations (Brown, 2025; Musthafa et al., 2024).

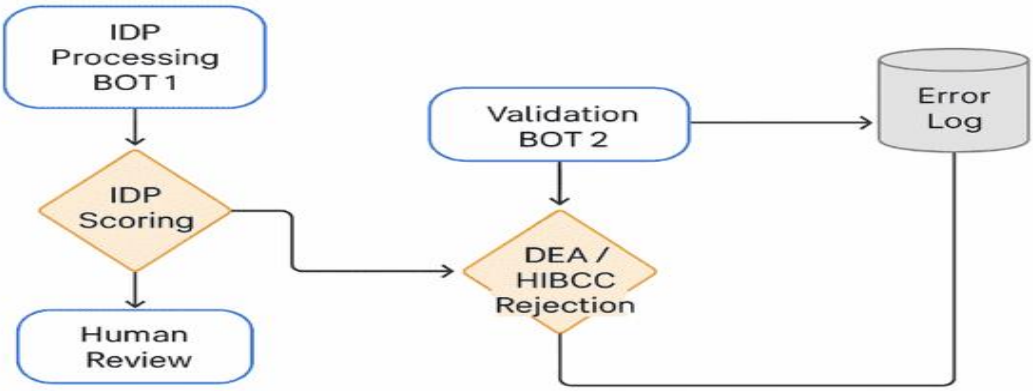


Figure 7: Error Control and Compliance Workflow in CRT Automation

5.3 Scalability and Audit Readiness

The CRT automation framework is eminent for its inherent scalability. When an organization builds out its distribution system and massive client onboarding places an order for such organization, the crunch comes in processing large volumes without an equal increase in manning. The present automation framework is modular and can be expanded horizontally by deploying more UiPath bots or vertically by integrating additional sources of validation and document types (Majeed, 2024; Ravindran & Balachandran, 2025).

Scalability also enables segregation by geography and business units. Application of metadata to each form and automation instance, including region, business line, and product type, allows the framework to provide real-time reports with performance indicators segmented according to key strategic zones. These insights can feed back into operational decision-making to understand which teams, regions, or customers need to be considered for policy amendments or training (Guevara et al., 2024).

In terms of audit readiness, multiple layers of evidentiary capture make this system ideal. UiPath bots record each transaction, including API calls, validation decisions, and decisions for fallback paths. The logs are archived and can be pulled easily for internal audits or checks of regulatory compliance (Dangol, 2025). Power BI dashboards built on SharePoint and UiPath logs allow compliance officers to keep real-time performance tracking of KPIs such as exception rate, response time, and validation failures over time with even greater weight, the premise for automated CRT is that it treats each customer record under uniform rules, ranking bias or inconsistency allowed by manual processing. Enforcement of such uniformity ensures legal defensibility and compliance with internal standard operating procedures and regulatory requirements (Yeoh, 2024; Hartikainen, 2024).

Table 5: Scalability Parameters of CRT Automation System

Parameter	Capacity before Automation	Post-Automation Capacity	Uptime (%)	Audit Log Availability
Forms Processed per Day	20 to 30	200 to 250	99.8	Full Traceability through SharePoint
Validation Throughput (DEA/HIN)	10 to 15 per hour	100+/hour	99.9	Exportable with Timestamped Entries
Processing Time per Form	45 to 60 minutes	3 to 5 minutes	99.7	Real-time Logs of Bot Activity
Exception Handling Resolution	24 to 48 hours	< 4 hours	99.5	Integratable to Power BI Alerts
Multi-Bot Parallel Execution	Not Applicable	5 to 10 concurrent bots	99.9	Logged per Bot Session ID

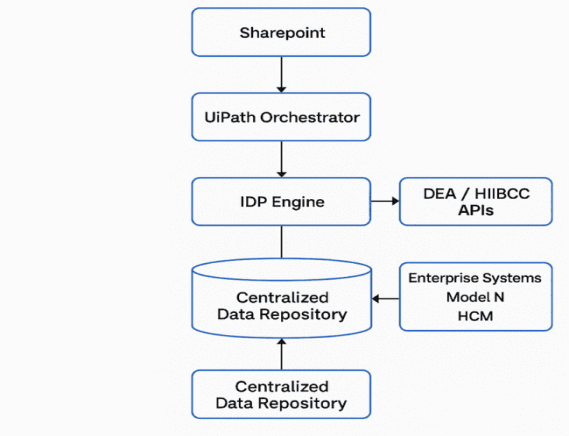
VI. SYSTEM ARCHITECTURE

6.1 Interaction of Components Across the Automation Framework

The Customer Record Tracking (CRT) automation process runs on a system architecture that is a tightly knit framework involving document management, robot layers of automation, artificial intelligence engines, regulatory web services, and enterprise data systems. Hence, the architecture is meant to guarantee high levels of reliability, traceability, and scalability, while conforming to regulatory requirements and internal compliance (Guevara et al., 2024; Majeed, 2024).

The architecture commences with SharePoint as the document repository and workflow automation launching pad. As the DEC form is uploaded into the SharePoint library, a record is created with timestamped metadata, including document source, submission time, and uploader identities. This event then triggers the execution of BOT 1, having been set up in UiPath Orchestrator to perform a set of workflow activities (Patil & Jadhav, 2023; Matt, 2025).

BOT 1 then feeds the data to the Intelligent Document Processing (IDP) engine for automated extraction. IDP uses optical character recognition (OCR) for recognizing text and natural language processing (NLP) models for semantic interpretation of the form fields. This interaction layer is configured with pre-trained document templates that could distinguish key-value pairs and be configured for retraining as time goes by to improve its accuracy (Musthafa et al., 2024). After extraction, the data is temporarily stored within SharePoint lists and tagged with a transaction ID for process tracking and error monitoring.



CRT System Architecture Overview

6.2 Data Flow Between SharePoint (S), IDP, External Portals, and Internal Systems

Suggested are classifiable and sequential dataflow architectures. When the IDP is done performing the form parsing, requiring shipping of the processed data via encrypted channel to BOT 2, which, with the external portals, carries out validation workflows comprising of two big integrations: into the DEA Registry and the HIBCC Directory (Mahe & Kapnisi, 2025; Lai, 2025).

BOT 2 attaches the relevant metadata (e.g., DEA number, business name, license expiration date) and then sends the information via secured REST API calls to validations. A response ("Valid," "Expired," or "Mismatch") in the form of JSON is logged into SharePoint, along with the link to the original form transaction ID, thereby providing both a real-time compliance check and full traceability of automated-made decisions (Brown, 2025; Oruç & Dupin, 2025).

After external validations are finalized successfully, the CRT system moves onto the data mapping phase within the internal systems. In this phase, the Model N and Human Capital Management (HCM) systems are accessed; and the architecture allows the BOT 2 to create new records or update existing ones. From inside Model N, bot aligns customers with pricing tiers, contract types, and sales hierarchies; and from inside HCM, it maps to corporate regions and reporting structure (Gloria et al., 2025). Back into SharePoint go the two systems, which keeps logs of every system-level interaction, inclusive of timestamps, change logs, and validation snapshots. There are monitors or dashboards for data flow, which monitor the latency, throughput, and failure rate over each integration node (Shirley & Kavin, 2024; Takács & Prorok, 2024).

Table 6: CRT System Integration Touchpoints

System Name	Data Consumed	Data Output	Interaction	Failure Handling Logic
SharePoint	Upload DEC form, metadata	Document status tags, and audit trail	UI/API	Retry 3 times, then escalate to admin queue
IDP Engine (UiPath)	PDF/Image-based forms	Structured key-value metadata	Embedded component	Flag low confidence fields for review
DEA Portal	DEA Registration Number, Entity Name	Init-License validity and status, expiry date	REST API	Retry on timeout, notify compliance
HIBCC Directory	HIN, Organization Name	Classification Type, Confirmation of Status	Web API	Suspend and alert if HIN cannot be found
Model N	Customer metadata, eligibility tags	Record create/update, pricing-alignment	API	Queue for retry, rollback on error
HCM (HR/Finance)	Entity name, region, hierarchy tags	Internal mapping confirmation	UI/API	Manual push if hierarchy mismatch

6.3 Security Layers, Scalability, and Failover Mechanisms

Regulatory sensitivity of CRT data starts with, for instance, customer DEA information and institutional identifier data flow—in accordance with this sensitive data is encrypted under Transport Layer Security (TLS) protocol. UiPath bots access APIs through OAuth tokens or managed identity credentials, so there exist no hardcoded passwords in this. Role-Based Access Controls (RBAC) are enforced throughout each layer of architecture, thereby ensuring that data is altered only by authorized users or processes (Guevara et al., 2024; Hartikainen, 2024).

Scalability is built into the design of the workflow through UiPath Orchestrator, which is capable of dynamically allotting bots depending on workload. At heavy submission cycles, i.e., beginning of a fiscal year or contract renewal season, bots are spun up on their own to deal with the increased data volume. Hence, without any manual operation, such elasticity keeps the system responsive (Yeoh, 2024; Ravindran & Balachandran, 2025). At critical interaction points, failover measures are set up. In the case where the DEA portal validation could not go through due to an external service outage, then BOT 2 places the process in deferment and retries it at a particular interval, along with alerting the stakeholders via automated SharePoint alerts. Records are stored temporarily in the holding queue if Model N becomes unresponsive with retry logic already built into the automation flow (Matt, 2025; Gloria et al., 2025).

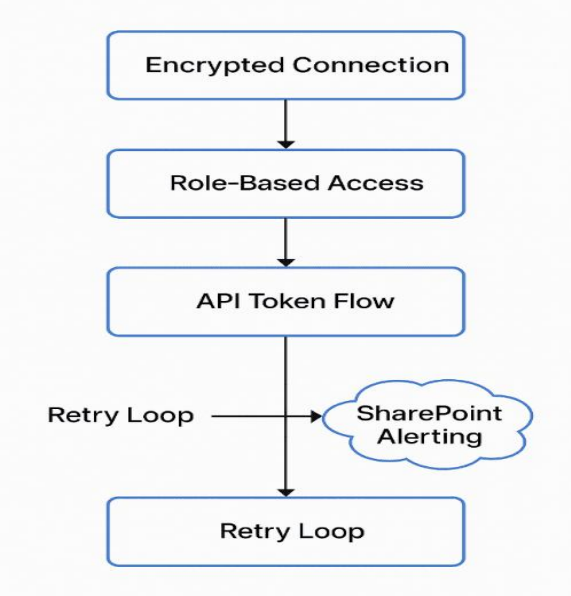


Figure 9: Security and Failover Layers in CRT System

VII. LIMITATIONS & OUT-OF-SCOPE

7.1 Handling Multi-HIN and Multi-DEA Records Within a Single Submission

A major limitation with the current CRT automation framework is that it cannot support multiple HINs or

DEA registrations embedded within a single PDF form. Unless the IDP engine employed by BOT 1 is optimized for structured document parsing, the current design imposes constraints on extracting and validating only one main entity per form submission (Shirley & Kavin, 2024; Patil & Jadhav, 2023).

Forms consisting of nested tables or having multi-entity declarations—for example, facilities affiliated with an umbrella organization having more than one DEA registration, or perhaps satellite clinics bearing separate HINs—are difficult to process. BOT 1 operations identify a single set of key-value pairs, whereas extracting multiple, overlapping sets of data records from a single form race with finesse into field collisions or misclassifications or even overwriting of one record over another (Musthafa et al., 2024).

While multi-entity detection might be supported with further enhancement in OSR embedded with deep-learning, this feature is still in the experimental phase and not yet deployed for production. Additionally, further validation and record creation logic routed for such complicated documents would likely arise violations when adapted for single-entity forms. risks and downstream integration failures (Mahe & Kapnisi, 2025; Majeed, 2024).

Therefore, the users have to submit separate forms for each entity or facility that is holding a unique DEA or HIN registration. Any failure to do so results in an automation stop, followed by alert notifications on SharePoint that flag the form for manual review. Though manageable in low-volume scenarios, this limitation acts as a bottleneck in large institutional networks or distributor chains, where grouped records form common practice (Lai, 2025).

Table 7: Multi-HIN/DEA Submission's Influence on the Automation Accuracy

Document Type	Number of Records Detected	Processing Accuracy	BOT Result	Manual Escalation Rate (%)
Standard Single-Entity DEC Form	1	98–99	Successful	2
DEC with Two DEA Registrations	2	65–70	Partial Extraction	30
DEC with Multi-HIN Entries		55–60	BOT Failed, Flagged for Review	45
Scanned DEC with Multi Entities	2-4	40–50	Extraction Incomplete	60
Docusign DEC with Multi Identifiers	2+	<40	Skipped by BOT 1	75

7.2 Docusign-Based Form Exceptions and Incompatibilities

A second major restriction concerns documents that are processed through Docusign, an e-signature and verification platform that has widely been used in healthcare for compliance documentation. While duly authenticated by Docusign and the chain-of-custody, the PDF structure might conflict with the needs of an IDP system and BOT parsing.

Docusign PDFs commonly embed their form fields within vector layers or image objects instead of plain text-based metadata, thereby presenting issues for typical OCR tools attempting to capture the data. BOT1, which depends on an IDP module to locate textual data to link to field names and field relationships, confronts difficulties in handling Docusign files that are image-heavy or encrypted and lack semantic tagging or adequately accessible layer structures (Dangol, 2025; Guevara et al., 2024).

Consequently, automation failures are very much likely to occur during the extraction phase for forms originating or recently modified from Docusign. Some such problems include field boundary definitions that are missing, or the duplication of metadata, or entire sections rendered unreadable because of digitally watermarked overlays. Very importantly, the Docusign

proprietary structuring resists any form of modification thus flattening without compromising the legal integrity of the document (Gloria et al., 2025).

Due to such great complexity, Docusign forms must be first manually altered or otherwise routed through another document transformation utility prior to upload into Sharepoint for standard automation processing. This gives the human operator another step and thus slightly reduces the overall automation efficiency in the environment. 2.2.3 Exceptions from Submissions via Docusign (Takács & Prorok, 2024).

Although roadmap discussions around the UiPath and Docusign integration APIs have started, native compatibility within BOT 1 and Docusign file architecture is still a matter of R&D. Hence, Docusign-based exceptions constitute a known limitation of the current CRT automation design (Yeoh, 2024; Matt, 2025).

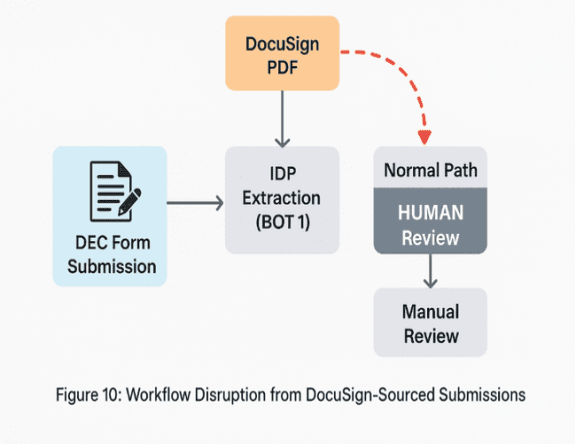


Figure 10: Workflow Disruption from DocuSign-Sourced Submissions

This visual should display how Docusign-pdf submissions deviate away from the normal BOT 1 extraction path to manual review.

7.3 White-Box Tasks and Business Logic Excluded from the Automation Scope

White-box tasks imply that a certain process requires contextual judgment, multi-system decision-making, or dynamic reasoning, and these are beyond current rule-based automation or supervised machine learning algorithms. In the CRT automation context, certain high-risk validation, legal disputes, or atypical onboarding scenarios have been considered out of the scope of robotic processing.

For example, it cannot be reliably automated to validate DEA registration anomalies that involve name mismatches caused by mergers, DBA (Doing Business As), or expired licenses that are legally subject to dispute. The scenarios require subjective judgments, escalation to legal teams, or cross-validation on multiple human-reviewed systems (Brown, 2025; Oruç & Dupin, 2025). Lastly, conditional pricing based on internal negotiating, dynamic contracting, and rules specific to a particular market lies outside BOT 2 integration with Model N. These sorts of arrangements mostly require human pricing committees or sales executive approvals and thus are classified as white-box (Mahe & Kapnisi, 2025).

Similarly, process escalations arising out of regulatory disputes, legal flagging, or audit-triggered customer data freezes outside the scope. While automation logs and audit trails are captured to aid these reviews, at least at this stage, the decisions are manually taken for reasons of legal defensibility and compliance with internal control frameworks (Guevara et al., 2024).

The exclusions are by design, supporting a hybrid automation pattern, automating repetitive, high-volume activities while escalating strategic exceptions to human judgment. Trying to automate such a decision-rich exercise would not only introduce technical errors but may well land the organization on the wrong side of legal and ethical boundaries (Majeed, 2024; Ravindran & Balachandran, 2025).

Table 8: CRT White-Box Task Exclusions and Justifications

Task Type	Reason for Exclusion	Risks if Automated	Manual Resolution Path	Compliance Impact
DEA Name Mismatch Resolution	Requires legal/contextual interpretation	False record rejection or approval	Escalation to Legal/Compliance for review	High – May lead to regulatory misalignment
Disputed License Status Investigation	Depends on case-specific correspondence with agencies	Inaccurate status update	Verified manually via DEA or HIBCC communication	High – Affects controlled substance handling

Pricing Tier Assignment in Special Contracts	Varies by deal, customer tier, and region	Wrong pricing or contract eligibility	Handled by Pricing/Finance Committee	Medium – Revenue leakage risk
M&A Entity Record Consolidation	Complex judgment needed for merging legacy data	Duplicate or corrupted master records	Data Governance team to approve and consolidate	High – Impacts data integrity across systems
Audit-Triggered Record Freezes	Requires cross-departmental coordination	Unauthorized edits to frozen profiles	Freeze maintained until audit team lifts flag	High – Violates audit preservation policy

VIII. CONCLUSION

8.1 Summary of Automation Value in CRT Systems

The Customer Record Tracking (CRT) automation process is an innovative approach that transforms how healthcare organizations address regulatory and operational complexity in customer profile onboarding and maintenance. At the foundation, the shift towards automation of processes has, through technology solutions like UiPath, SharePoint, and Intelligent Document Processing (IDP), shown value in directly decreasing turnaround time, improving accuracy, and reinforcing compliance assurance (Majeed, 2024; Patil & Jadhav, 2023).

Over 500 hours saved every quarter, and with more than 7,100 hours estimated annually, the gains are not only evident and measurable but, in fact, also can be leveraged among several business units (Takács & Prorok, 2024). These have resulted from eliminating any form of repeated manual interference, allowing optimized document ingestion and classification along with automated cross-platform validations against the DEA and HIBCC portals. Further, the integration into internal systems such as Model N and Human Capital Management (HCM) folds out fewer challenges as record synchronization supports consistent pricing, hierarchy mapping, and contract eligibility tracking (Matt, 2025; Musthafa et al., 2024). By being able to have audit-ready logs, traceable validation paths, and exception-handling workflows, they offer and improve regulatory compliance and possibilities for controls concerning internal governance. Automation, too, has helped smooth the way for user experience, allowing for effortless submission of the forms and thus removing the administrative load off internal compliance teams to support a more agile operational mode (Mahe & Kapnisi, 2025; Guevara et al., 2024).

Technical limitations continue to exist, of course, including those with multi-DEA/HIN forms and against Docusign-type structures; the CRT framework effectively separates these exceptions so that the integrity of the larger process is not compromised. Human-in-the-loop verification methods also help keep a balance reflected in automation and accountability and give way to continuous learning and process improvement (Oruç & Dupin, 2025; Shirley & Kavin, 2024).

8.2 Strategic Benefits: Efficiency, Scalability, and Data Quality

The CRT automation venture may have strategic implications well beyond the sphere of task automation. From the point of view of efficiency, the system acts as an enabler for near real-time decision-making by condensing cycle time from several hours to minutes. This ability to respond fast to customers' requests, get contracts activated in no time, and review prices fast is an advantage to healthcare distributors and are all important factors on which to compete (Lai, 2025; Brown, 2025). From a scalability standpoint, the proliferation of the modular design through various business avenues, locations, or new product lines could be ensured there. The system architecting has been embedded with the capacity to add new document types, integrate new validation portals, or expand the automation layer to adjacent workflows, such as rebate processing or license renewal (Ravindran & Balachandran, 2025; Gloria et al., 2025). This way, extensibility is achieved, meaning that the CRT automation initiative remains not a set-in-stone solution but rather a moving foundation for general digital transformation.

Other aspects that see improvement are data quality. Before automation, customer records were often inconsistent, duplicated, and suffered from undue

delays when sensitive regulatory fields required updating. By means of automated validations, schema enforcement, and centralized data logging, the CRT now supports enhanced data quality and compliance reliability (Yeoh, 2024; Hartikainen, 2024). The enhancements pave the way for improved analysis, transparent audit trails, and better reporting of pricing and sales functions.

Lastly, it can be inferred that CRT automation has positioned the organization for long-term success. This improves operational resilience, mitigates compliance risks, and is overarched by an enterprise-wide digital strategy. As regulatory landscapes evolve and data volumes grow, CRT automation will thus become one of the necessary facilitators toward sustainable, scalable, and secured governance of customer relations. (Matt, 2025; Majeed, 2024).

REFERENCES

[1] Affandi, N. A. M., Zahibi, M. F. H. M., & Bakar, N. A. (2025). Driving digital transformation in healthcare. *Journal of Information and Knowledge Management*. https://journal.uitm.edu.my/ojs/index.php/JIKM/article/view/6156

[2] Brown, J. (2025). Cognitive automation in regulatory compliance workflows. *IEEE Conference Proceedings*.

[3] Dangol, R. (2025). Process consolidation via automation. *ProQuest Dissertations Publishing*.

[4] Galindo Vásconez, D. F. (2024). Implementación de un sistema informático para gestión de citas. *Repositorio UTA*. https://repositorio.uta.edu.ec/bitstreams/ae0993d1-d362-47d4-b46f-7ba553018886/download

[5] Gloria, E., Diaz, K., & Celis, V. A. (2025). Leveraging web technologies in public health workflows. *IEEE Xplore*. https://ieeexplore.ieee.org/document/10879648

[6] Guevara, A., Ballena, H., & Mansilla-Lopez, J. (2024). Managing service payments and automation risks in compliance-sensitive workflows. In *Information, Innovation & Impact*. Springer.

[7] Hartikainen, H. (2024). Improving compliance in medical device software and healthcare automation. *Aalto University School of Science*.

[8] Lai, K. W. (2025). Automation in phototherapy and digital audit readiness. *Springer Nature*.

[9] Mahe, L., & Kapnisi, K. (2025). Adoption of automated bioprocessing and compliance platforms in CGMP. *Cytotherapy*, Elsevier.

[10] Majeed, M. (2024). Artificial intelligence and business compliance strategy. *Bentham Science Publishers*.

[11] Matt, H. (2025). Machine learning and robotic process automation: Efficiency benchmarks. *ResearchGate*. https://www.researchgate.net/publication/391653435

[12] Musthafa, A. S., & Bhuvaneshwari, S. (2024). Enhancing data accuracy with machine learning in healthcare systems. *IEEE International Conference on Information Systems*.

[13] Oruç, M. F., & Dupin, M. (2025). Enhancing auditability through API-driven compliance workflows. *Applied Marketing Analytics*.

[14] Patil, V., & Jadhav, P. (2023). RPA and IDP integration for intelligent document classification. *ResearchGate*.

[15] Patil, Y. H., & Patil, R. Y. (2024). AI-powered platforms in health compliance automation. *Google Books*.

[16] Ravindran, S., & Balachandran, G. B. (2025). IoT and automation for data governance and risk reduction. *Journal of Industrial Information*.

[17] Shirley, C. P., & Kavin, S. (2024). Intelligent OCR and pattern recognition for medicine packaging and customer forms. *IEEE Conference on AI in Medicine*.

[18] Shukla, A. (2025). Real-time automation in embedded healthcare compliance. *HS Offenburg University Publications*.

[19] Takács, I., & Prorok, M. (2024). Business transformation through RPA in regulated industries. *IEEE Jubilee Conference on Business Informatics*.

[20] Yeoh, P. S. Q. (2024). Centralizing healthcare databases and ensuring data integrity. *Springer Publications*.