**Business Requirements Document (BRD)**

**Project Title:** Auditing Process Automation  
**Document Version:**   
**Date:**

**Table of Contents**

[1.1 Project Overview 1](#_Toc3300264)

[1.2 Project Scope 1](#_Toc984355531)

[1.3 Stakeholders 1](#_Toc544756994)

[2. Business Objectives 1](#_Toc902516446)

[3. Current Process Overview 1](#_Toc152306088)

[4.Proposed Solution 1](#_Toc1909636429)

[4.1 Proposed Solution 2](#_Toc839041395)

[5. Requirements 2](#_Toc1364577818)

[5.1 Functional Requirements 2](#_Toc938266663)

[5.2 Non-functional Requirements 3](#_Toc509151368)

[6. Assumptions and Constraints 3](#_Toc1222420832)

[6.1 Assumptions: 3](#_Toc1106089040)

[6.2 Constraints: 3](#_Toc559753128)

[7. Risks and Mitigation 3](#_Toc692825175)

[8. Project Timeline 4](#_Toc998891819)

[9. Approval 4](#_Toc394191145)

[10. Revision History 4](#_Toc2134461422)

[11. Conclusion 4](#_Toc459982937)

1.1 Project Overview  
The auditing process for MSN (Medical Non-Standard) plans renewal requires an automated approach to ensure accuracy and consistency between the CIRRUS database and the sources of truth, so that the correct information from CIRRUS is used to generate the ID Cards.

# 1.2 Project Scope

The auditing process will focus on comparing MSN plan details extracted from CIRRUS with the sources of truth. There approximately 371 plans with total of 31 fields to audit, and the sources of truth include a BARTRACK Report (for Oxford Book of Business) and other reports for other books of business. The fields to audit include {}. Consideration will be given to managing multiple benefit plans within a single group during the auditing process. The project attempts to extract relevant information from the CIRRUS database using queries based on the relevant fields needed for filtering the data. Additionally, the aim is to gather relevant data from all the sources of truth and arrange them in a consistent format, thereby addressing the challenge posed by the distorted and inconsistent information of the sources of truth.

# 1.3 Stakeholders

# 2. Business Objectives

The primary objective is to automate the auditing process to accurately align MSN plan details extracted from the CIRRUS production environment with the information provided in the sources of truth. This aims to minimize errors and discrepancies in membership ID cards issued during plan renewals. This also saves time and reduces the scope of human error while manually auditing the required fields.

# 3. Current Process Overview

## 3.1 Manual Audit Process:

1. Access USP Plans Active and termed database link from CIRRUS: This is an extract from CIRRUS which contain details of active member groups, benefit plans and effective renewal dates.

2. Filter the MNS (Medical Non Standard) plans: From the Active report, the user has to filter only the MNS plan data. MNS plans are specifically tailored/customized to the needs of each member group.

3. Download the active report in CSV format: This will give us the names of groups which we need to audit.

4. Open the downloaded file and filter on Book of Business: Notate what all Book of Businesses we have

5. Filter on Plan ID column and input MNS in the search bar

6. Filter on CONTRACT OPT EFF DATE column and uncheck the Select All checkbox: Select appropriate month check box which you want to audit

7. Select and copy all the values in the Plan ID column and paste them in a new sheet

8. Click on the Data tab and ‘Remove duplicates’

9. Verify SITUS STATE column and notate the State values (example: NY, NJ, CT)

10. Click on the 2nd tab in the excel file where the benefit plans are pasted

11. Input the corresponding State values next to each benefit plan: Now we know what plans need to be audited and for what States

12. Access CIRRUS benefit configuration sharepoint: In the CIRRUS benefit configuration sharepoint, scroll to Non Standard Plan Documents section and click on the appropriate BARTrack icon (currently there are CT BARTrack, NJ BARTrack, NY BARTrack)

13. Click Open in Desktop from the pop up box if applicable

14. Toggle back to the active detail file where we have the list of plans and state

## 3.2 Challenges faced

The inconsistent formats of sources of truth presents a challenge. The manual comparison between the CIRRUS data and the sources of truth makes it inevitable to cause potential errors and inconsistencies. The inconsistent storage format of sources of truth adds complexity to the manual matching process.

The proposed automation aims to streamline this process, ensuring accuracy and efficiency in extracting and comparing data.

# 4.Proposed Solution

## 4.1 Proposed Solution

Our solution is a 2-step approach, where we first automate the auditing of those plans which could be successfully standardised and matched with CIRRUS, and in the next step we create a GUI application for those plans whose auditing could not be automated.

1. Standardization of sources of truth:

* Convert Excel workbooks to PDF format to ensure uniformity and compatibility.
* Utilize DocumentAI, a powerful document processing tool, for structured text extraction, enabling efficient data extraction from the reports.
* Employ Gemini AI prompt engineering to develop tailored prompts for extracting relevant field values, ensuring accuracy and completeness.
* Extract embedded files and store them in a designated folder for easy access and reference.
* Aggregate the embedded data into a single PDF and repeat step 2 and 3 for further processing, simplifying data handling and management.
* Store all the extracted information into an Excel Workbook.
* {BARTRACK is already standardised, need to standardise the other sources of truth}

2. Standarization of sources of truth

* Use table transformation and manipulation techniques to bring all the required data from BARTRACK reports into a single dataframe
* Use a custom made LayoutLM model to extract the relevant fields and values from images/forms
* Store all the extracted data into a dataframe
* Combine all sources of truth by appending the above dataframe with the one generated from the BARTRACK report
* Convert the combined dataframe into a csv/excel file

3. Data Retrieval from CIRRUS Database:

* Use SQL queries tailored to extract relevant information for active MSN and standard plans from the CIRRUS Database.
* Ensure optimized queries for efficient data retrieval, minimizing processing time and resource utilization.

4. Data Comparison and Error Detection:

* Employ advanced Natural Language Processing (NLP) techniques or text matching algorithms for accurate comparison of field values.
* Compare extracted field values from CIRRUS with standardized BARTRACK reports, which serve as the Source of Truth.
* Flag any inconsistencies or discrepancies detected during the comparison process as errors for further investigation and resolution.
* Identify those plans on which text matching techniques failed

5. Auditing using Tkinter GUI:

{}

6. Automation and Efficiency:

* Streamline the end-to-end process through automation to enhance accuracy and efficiency.
* Minimize manual intervention, reducing the risk of errors and ensuring consistent results.

4.2 **Key Features and Functionality**

4.3 **Technology Stack**

# 5. Requirements

## 5.1 Functional Requirements

**Requirement 1:**

* + **Description:**
  + **Acceptance Criteria:**

**Requirement 2:**

* + **Description:**
  + **Acceptance Criteria:**

## 5.2 Non-functional Requirements

**Requirement 1:**

* + **Description:**
  + **Acceptance Criteria:**

# 6. Assumptions and Constraints

## 6.1 Assumptions:

* The CIRRUS system is the primary source of data for member groups, benefit plans, and renewal dates. The sources of truth, compiled from data entered by the sales team, and considered to be reliable for auditing purposes. Both the CIRRUS system and Bartrack reports are easily available.
* There is no fixed timeline for running the automation for auditing process. Typically the renewals are done on a specific month each year, but there might be a need for daily automation during peak renewal periods.
* 90% of the sources of truth lie in the BARTRACK report, which contains plans present in the Oxford Book of Business. The other books of business such as TUFT, AIMS, etc. Have their own sources of truth.
* The BARTRACK Report is an Excel Workbook consisting of standardized tables where each worksheet contains plans for each group/client.
* The format of documents submitted by sales into the Bartrack system for each corresponding group remains consistent across all renewal periods.

## 6.2 Constraints:

* The unstructured and inconsistent format of the sources of truth may pose challenges in standardizing and bringing together all the information under one roof.
* The dynamic nature of embedded files within BARTrack reports necessitates careful handling and validation to ensure accuracy.
* The large volume of groups undergoing audit on a regular basis may impact the runtime of automation, requiring optimization for efficiency.
* Since the algorithms being used always have some level of error rate, manual intervention may be necessary to verify discrepancies and ensure 100% accuracy in the audit process.

# 7. Risks and Mitigation

* Data Consistency and Accuracy:
  + Risk: The reliance on consistent CIRRUS reports and inconsistent sources of truth, introduces the risk of error while matching the datasets, potentially leading to inaccurate audit results.
  + Mitigation: Implement robust data validation procedures to ensure consistency and accuracy between CIRRUS and BARTrack data, such as cross-referencing key fields and conducting regular reconciliation checks.
* Structural Limitations to sources of truth other than BARTrack Reports
  + Risk: Unstructured sources of truth (especially images) poses a risk to the extraction and detection of key fields and completeness of the audit process and may hinder the ability to perform accurate comparisons.
  + Mitigation: Use a deep learning based model to annotate and label the keys and values in the images, facilitating a smooth text extraction process
* Dynamic Nature of Embedded Files
  + Risk: The presence of embedded files within BARTrack reports introduces complexity while extracting data and increases the likelihood of errors during data extraction and validation, potentially impacting the accuracy of the audit process.
  + Mitigation: Develop automated processes for extracting and validating data from BARTrack reports, leveraging OCR and ML algorithms to handle embedded files and minimize the risk of manual errors.
* Manual Intervention for Verification
  + Risk: Since ML algorithms are never error free, manual intervention to verify and audit the undetected groups/fields poses a risk of human error and may slow down the audit process, especially during peak renewal periods when daily automation is required.
  + Mitigation: Implement standardized workflows and checklists for manual verification of discrepancies, ensuring that the process is efficient and consistent across all audits, with clear escalation procedures for resolving discrepancies promptly. For example- provide a clear dashboard of the fields/groups the automation couldn’t audit and provide all additional details to track the same and reduce manual effort.

# 8. Project Timeline

# 9. Approval

This BRD requires approval from relevant stakeholders before proceeding with implementation.

# 10. Revision History

Any revisions or updates to the BRD will be documented and maintained for future reference.

# 11. Conclusion