Functional Document – HIPAA 834 EDI File Processor with AI/ML Roadmap

# 1. Objective

The goal of this project is to build a scalable and automated file processor for HIPAA 834 EDI and related files.  
- Process incoming files from state-wise shared network directories.  
- Apply HIPAA validation (existing API/logic).  
- Generate Excel reports with error details and timestamps.  
- Organize processed files into Archive folders automatically.  
- Introduce AI/ML (ML.NET) to detect anomalies (e.g., unusual file sizes, missing loops, unexpected patterns).  
- Provide an optional dashboard for monitoring and insights.

# 2. Functional Flow

1. Configuration-driven: A JSON config defines each state’s directory and processing frequency (Daily / Monthly). The processor reads this config to decide which folders to scan.
2. File Detection & Copying: Files are scanned in the shared NAS directories. Files are copied to a local working directory for processing.
3. HIPAA Validation: Each file is passed to the HIPAA Validator API. The API returns error details, which are saved in Excel reports. Reports are timestamped and stored in the Report folder (within the state directory).
4. Archiving: Once processed, files are automatically moved to Archive/yyyy-MM-dd/ subfolders.
5. AI/ML Integration (Phase 1): Collect metadata (file size, processing time, error counts). Use ML.NET anomaly detection to highlight unusual trends. Save results in a separate Excel anomaly report with charts.

# 3. Key Features

✅ Automated folder scanning based on schedule (daily/monthly).

✅ Configurable (no hardcoding of paths).

✅ HIPAA-compliant error reporting.

✅ Self-organizing Reports and Archive folders.

✅ AI/ML insights to detect data quality anomalies.

✅ Optional Dashboard for centralized monitoring.

# 4. Technology Stack

- .NET Framework 4.7.2 (Console Application)  
- App.config / JSON for configuration  
- ML.NET for anomaly detection  
- ClosedXML for Excel reporting with charts  
- (Future) ASP.NET MVC / Web API + React or Power BI for dashboard

# 5. Benefits

• Reduces manual intervention in EDI file processing.

• Ensures compliance with HIPAA by validating each file.

• Provides traceability with timestamped reports.

• Improves data quality monitoring with AI/ML.

• Scalable for new states, frequencies, or formats (.EDI, .P, etc.).

• Provides real-time visibility via dashboard (future).

# 6. Future Enhancements

→ Dashboard (web or BI-based) for centralized monitoring:  
 - Real-time processing status (Success/Failure).  
 - Error count trends and anomaly charts.  
 - Drill-down by state, by file, by error type.

→ Predictive Analytics: forecast spikes in errors or file sizes.

→ Self-Healing: automated retries for failed file validations.