Maryland Department of Human Services
Office of Benefits Access
AI-Powered Predictive Error Reduction System: Technical & Strategic Blueprint
September 15, 2025
Version 1.0 - Proposal Draft for Arnold Ventures

Executive Summary

This document provides the technical and strategic blueprint for an AI-Powered Predictive Error Reduction System. The proposed system is a key component of our strategy to enhance the integrity and efficiency of the Supplemental Nutrition Assistance Program (SNAP). By leveraging advanced analytics and existing state infrastructure, this system will proactively identify cases with a high risk of payment errors, allowing for preventive intervention. This blueprint details the project's strategic goals, technical architecture, data modeling approach, caseworker integration, and the essential governance framework required for a successful multi-state pilot.

1.0 Strategic Purpose

The system's primary purpose is to **directly reduce SNAP payment error rates** by shifting the agency from a reactive "pay and chase" model to a proactive "predict and prevent" strategy. By using predictive analytics, the system will identify active SNAP cases that have a high statistical risk of future income-related or procedural errors *before* they occur. This approach enhances program integrity and ensures fiscal responsibility by mitigating the risk of direct financial penalties tied to a high Payment Error Rate (PER).

2.0 Technical Architecture

The system will be designed as a **modular component** that integrates directly into **Maryland's MD-THINK platform**. This strategy leverages the state's prior technological investments and avoids the high cost and risk of building a new system from the ground up.

- **Platform Integration**: The system will be built as an independent microservice that communicates with the core MD-THINK platform via a secure, well-defined Application

Programming Interface (API), aligning with the platform's modular architecture.

- **Cloud-Native Foundation**: It will be built on the secure, scalable cloud infrastructure of Amazon Web Services (AWS), which provides the foundation for the MD-THINK platform.
- **Data Source**: The engine will connect to the platform's **shared data repository**, which consolidates information from multiple health and human services programs to create a unified view of a client's circumstances.

3.0 Data Modeling

A supervised machine learning model will be developed to analyze anonymized data patterns and generate a real-time "error risk score" for each active SNAP case.

- **Model**: A logistic regression or gradient boosting model will be trained on historical, anonymized quality control (QC) data to identify the key factors that predict a payment error.
- **Data Sources**: The model will analyze anonymized, aggregated data patterns from the platform's shared data repository, including:
  - **Wage Data**: Changes in income detected in **quarterly W-2 data** from the state's Department of Labor.
  - **Household Data**: Changes in household composition sourced from vital statistics or other linked programs.
  - **Case Data**: A history of frequent income changes, past procedural errors, or inconsistencies in self-reported information.

4.0 Caseworker Integration

The system is designed to **augment caseworker expertise**, not replace it. It will be seamlessly integrated into the daily workflow to help staff prioritize their efforts.

- **Risk Dashboard**: An intuitive dashboard will display a prioritized list of high-risk cases, allowing caseworkers and supervisors to focus their attention where it's needed most.
- **Intelligent Nudges**: For each high-risk case, the system will provide caseworkers with **"intelligent nudges"** and pre-analyzed summaries that highlight the specific risk factors.
- **Explainable AI (XAI)**: Every alert will include a clear, plain-language explanation of *why* a case was flagged (e.g., "New W-2 data does not match reported income"), building caseworker trust and ensuring transparency.

5.0 Governance & Ethics

A robust governance and ethics framework is crucial for building trust and ensuring the system is used responsibly.

- **Human Oversight**: A multi-stakeholder **"Citizen Data Stewardship Council,"** composed of independent privacy advocates, civil liberties experts, and citizen representatives, will provide ongoing oversight on data use and ethical protocols.
- **Bias Mitigation**: The model's logic and training data will be rigorously audited for demographic or socioeconomic bias before deployment and monitored on an ongoing basis. All final case decisions will continue to be made by human caseworkers.
- **Transparency**: The system will be designed for transparency, allowing program administrators and auditors to understand the logic behind its risk scores, leveraging Explainable AI (XAI) principles.
- **Data Security**: All data handling will comply with a comprehensive **Data Governance Framework** that ensures adherence to stringent security regulations like IRS 1075 and HIPAA.