# Strategic Roadmap: Tower Loan AI Fraud Co-Pilot

**Engagement Model:** Staff Augmentation & Enterprise Scale

**Effective Start Date:** January 7, 2026

**Prepared For:** Tower Loan & Internal Engineering Teams

## 1. Executive Summary

Following the successful Proof of Concept (POC), where we demonstrated the ability to detect high-risk anomalies and identify "un-lendable" profiles (including applicants with active lawsuits), we are transitioning to a **Staff Augmentation** engagement model.

This shift allows our engineering team to function as a direct extension of Tower Loan’s internal technology group. Unlike a fixed-scope project, this engagement focuses on continuous evolution, operational excellence, and the long-term maturity of Tower Loan’s AI infrastructure.

Our primary objective is to evolve the "Fraud Detector" from a static experiment into a dynamic, enterprise-grade **AI Co-Pilot**. This system will reside entirely within Tower Loan's AWS environment, protecting capital at the pre-approval stage by leveraging deeper historical data and real-time operational feedback.

## 2. Strategic Methodology: The Horizon Approach

As we move into a Staff Augmentation capability, we will execute against **Strategic Horizons** rather than rigid calendar deadlines. This allows us to adapt to technical discoveries (e.g., data quality in legacy tables) and shifting business priorities (e.g., accelerating Branch Fraud detection) without renegotiating scopes.

### **The Core Philosophy: "Ride-Along" Intelligence**

The AI will not replace human judgment; it will act as a high-precision "Ride-Along" partner, offering real-time risk scores and GenAI-powered explanations via API, accessed securely by your operational systems.

## 3. Horizon 1: Data Scaling & Foundational Deployment

**Theme:** *Re-platforming the POC to Production AWS & Preventing Overfitting.*

The POC proved the concept, but it was built on a limited dataset. To ensure the model performs in the real world without "overfitting," we must retrain it on a significantly larger volume of historical data securely within the Tower Loan AWS perimeter.

### **3.1 Secure Data Ingestion Strategy**

We will establish a secure, repeatable pipeline to fetch data from the on-premise/Colo SQL Servers (TLOS & OLL) into the Tower Loan AWS S3 environment.

* **Mechanism:** Implementation of a secure data fetch script (or AWS Direct Connect/VPN Tunnel validation) to pull data with **Read-Only** access.
* **Privacy & Security:** All PII (Personal Identifiable Information) will be hashed or encrypted *in-transit* and *at-rest* in S3.
* **Target Tables (Expansion):** We will pull bulk historical data from the specific tables utilized in the POC pipeline to ensure schema consistency:
  + **OLL DB:** ApplicantAddress, ApplicantDriversLicense, APPSQL\_OLL\_LoanApplication, FinancialInformation, PersonalInformation.
  + **TLOS DB:** Address, CreditReportV2 (various versions), EmploymentHistory, IdentificationInfo, Loan, LoanApplication, Payment, Person.

### **3.2 The "Headless" API Deployment**

We will deploy the inference engine as a "Headless" (UI-free) service to allow for testing without disrupting current operations.

* **Infrastructure:** Containerized inference code (Docker) running on AWS Lambda or SageMaker Serverless Inference.
* **Access Point:** An Amazon API Gateway secured via API Keys/IAM, allowing internal Tower Loan services to send a LoanApplicationID and receive a JSON fraud score + PDF Report URL.
* **GenAI Reporting:** Integration of Amazon Bedrock to generate the human-readable "Explanation Reports" that highlight *why* a loan was flagged (e.g., "Address mismatch combined with high-risk employment profile").

### **3.3 Model Recalibration & Validation (The "Compliance Check")**

Once the model is retrained on the larger dataset:

1. We will run a batch prediction on a "Blind Test Set" (recent applications the model hasn't seen).
2. We will submit the highest-risk profiles to the **Tower Loan Compliance/Audit Team** for verification (similar to the 4-case review in the POC).
3. **Refinement:** Based on feedback (e.g., "This wasn't fraud, but it was an uncollectible lawsuit"), we will tune the Business Rules and Class Weights.

## 4. Horizon 2: Operational Integration & The Feedback Loop

**Theme:** *Connecting the Brain to the Body.*

In this horizon, we integrate the "Headless" API into the actual Tower Loan loan origination workflow.

### **4.1 Defining the Trigger Point (Non-Deterministic)**

We will collaborate with your technical team to identify the optimal moment to trigger the API.

* **Scenario A (Pre-Approval):** Triggered immediately after the application is submitted but before the credit pull (saving credit bureau costs).
* **Scenario B (Pre-Funding):** Triggered before the final dispersal of funds as a final safety check.
* **Implementation:** We will assist in writing the "Hook" or trigger logic within your CRM or Loan Management System to call our AWS API Gateway.

### **4.2 The "AI Adjudicator" & Active Feedback**

A model is only as good as its latest data. We will implement a feedback mechanism to capture the human reviewer's decision.

* **The Workflow:** If the AI says "High Risk" but the Human says "Approve," the system must capture this **Override**.
* **Data Capture:** These decisions are stored in a new S3 bucket (/human-feedback-logs) to serve as "Gold Standard" labels for future training.

## 5. Horizon 3: MLOps Maturity & Scale

**Theme:** *Automation, Monitoring, and Expansion.*

As a Staff Augmentation partner, our goal is to move the system from "Manual Maintenance" to "Automated Health."

### **5.1 MLOps Maturity (Level 0** $\to$ **Level 1)**

* **Current State (Level 0):** Data extraction and model training are manual/script-driven processes managed by our engineers.
* **Target State (Level 1):** Automated Retraining Pipelines.
  + System automatically detects "Data Drift" (e.g., if applicant income distribution changes drastically).
  + Triggers a retraining pipeline in AWS SageMaker without human intervention.
  + Deploys the new model only if it beats the previous version's accuracy metrics (F1 Score, Recall).

### **5.2 Business & Technical Monitoring Dashboard**

We will implement dashboards (using CloudWatch or QuickSight) to visualize:

* **Technical Metrics:** API Latency, Error Rates, Model Drift.
* **Business Metrics:** Number of applications flagged, estimated capital protected, and distribution of identified personas (e.g., "Digital Ghost" vs. "High Friction Anomaly").

### **5.3 Expansion: Branch & Sales Fraud**

Once the OLL (Online) engine is stable, we will begin the **Discovery Phase** for Branch-based fraud.

* **Strategy:** Map the different data signals available in physical branches vs. online.
* **Action:** Replicate the MLOps pipeline structure to ingest Branch data, creating a unified Fraud Defense layer across all channels.

## 6. Technical Architecture & Security Standards

### **6.1 Cloud-Native Stack**

* **Compute:** AWS Lambda (for orchestration) & Amazon SageMaker (for training/inference).
* **Storage:** Amazon S3 (Tiered storage for Raw Data, Processed Data, and Model Artifacts).
* **GenAI:** Amazon Bedrock (Titan or Claude models) for text generation.
* **Orchestration:** AWS Step Functions to manage the dependency flow (Data to Train to Deploy).

### **6.2 Security & Governance**

* **Shared Ownership:** We operate within Tower Loan's AWS and DB environments, ensuring all intellectual property and data remain within the client's perimeter.
* **Encryption:** Every data pull involves encryption-at-rest and in-transit, with a strict policy of removing or masking PII before the modeling stage. All data operations utilize AWS KMS (Key Management Service).
* **Least Privilege:** IAM roles will be scoped strictly. The Staff Aug team will operate within the defined boundaries of the Tower Loan AWS account.
* **No "Black Boxes":** All code, model weights, and documentation are intellectual property of Tower Loan and reside in your repositories.
* **Flexible Capacity:** Our team’s focus can pivot between engineering, data science, and strategy based on the immediate needs of Tower Loan’s roadmap.

## 7. Conclusion

This roadmap represents a shift from "building a tool" to "engineering a capability." By leveraging the Staff Augmentation model, Tower Loan ensures that this AI Co-Pilot does not become a static legacy system but evolves into a dynamic, learning asset that adapts to new fraud vectors and protects the company’s capital in real-time.