

# Synthetic Data Generation for Fraud Detection Model Training

## Overview

Simthetic.AI is a technology company operating in the banking, financial services, and insurance (BFSI) industry, focused specifically on synthetic data generation to strengthen fraud detection capabilities without exposing sensitive real-world financial documents.

## Business Challenge

Training and stress-testing fraud detection models requires large volumes of diverse, realistic fraudulent documents to train against. Real examples such as invoices and purchase orders are sensitive and limited in supply and using them for training risks exposing confidential information.

- Fraud detection models need large volumes of diverse, realistic fraudulent documents to train against
- Real fraudulent documents are sensitive and limited in supply
- Rare fraud patterns and edge cases are hard to source from real data alone

## The Solution

A model was built to generate high-fidelity synthetic datasets of fraudulent documents, including invoices, purchase orders, and other document types. These synthetic datasets are used to train and stress-test fraud detection models, enabling rare fraud patterns to be simulated without exposing any real, sensitive data.

- Generates high-fidelity synthetic datasets of fraudulent documents, including invoices and purchase orders
- Used to train and stress-test fraud detection models
- Avoids exposing any real, sensitive data during model development

## Impact Delivered

The solution enabled Simthetic.AI to train and stress-test fraud detection models faster, using synthetic datasets that closely reflect real-world fraud patterns, while providing realistic coverage of rare, edge-case scenarios and ensuring no real, sensitive data was exposed during development.

- Faster training and stress-testing cycles for fraud detection models
- Realistic coverage of rare and edge-case fraud patterns
- No exposure of real, sensitive customer data during development
- Scalable fraud simulation supported across multiple document types

## Technology Stack

- High-fidelity synthetic data generation model
- Generative AI model for synthetic document creation
- Document-type templates (invoices, purchase orders, and others)
- Fraud pattern simulation engine

## Services Delivered

- Synthetic fraudulent document dataset generation
- Fraud detection model training and stress-testing support
- Privacy-preserving data generation with no real data exposure

Visit us online at <https://mothersontechnology.com/>

Motherson Technology Services Limited  
C- 26, Sector 62, NOIDA - 201309, U.P., India  
Email: [info.mts@motherson.com](mailto:info.mts@motherson.com)  
Tel: + 91 120-436-5555; + 91 120-436-5556

Proud to be part of.

