incode

INCODE FRONTIER AI LAB · TECHNICAL OVERVIEW

AI that evolves faster than fraud

Incode builds its own foundation models for identity, trained on billions of real verifications and stress-tested against the newest gen-AI attacks. This document details the models, the data engine behind them, and the governance that keeps them safe.

NIST #1 FACIAL RECOGNITION · IBETA LEVEL 2 LIVENESS · DHS RIVTD BENCHMARKS MET · FIDO FACE CERTIFIED

JULY 2026 · INCODE.COM/TECHNOLOGY

01 · FOUNDATION MODELS

Three models, one adaptive defense

VLM

Identity Vision-Language Model

Trained on global identity data, documents, templates, and synthetic fraud samples across 200+ regions. Analyzes visual and textual signals together to detect tampering, synthetics, deepfakes, and altered documents with high precision.

- 67% fewer errors in fake-ID detection
- 25% fewer errors in signature-fraud detection
- 0% APCER on a challenging liveness dataset
- +4.0% gain in document-tampering accuracy
- ~80% faster training cycles

LLM

Fraud Large Language Model

Trained on proprietary fraud datasets, identity metadata, behavioral sequences, device signals, and transactional flows. Interprets complex, multi-source patterns in real time to uncover hidden anomalies and intent.

- Anomaly and pattern detection across behavioral sequences
- Fraud-intent classification at enterprise scale
- Risk-signal extraction from messy, multi-source data

AGENTS

Reasoning Agents for Risk Intelligence

Trained on hundreds of signals and millions of labeled identity and fraud outcomes. Orchestrate outputs from the VLM, the LLM, device telemetry, and behavioral insight into a unified, context-aware risk assessment.

- Mexico: FAR improves by ≈47.6%, FRR by ≈60.3%
- Holistic risk decisions across multi-modal signals
- USA: FAR improves by ≈12.1%, FRR by ≈19.1%

02 · THE DATA ENGINE

What our models learn from

Models train on data from government databases, enterprise integrations, and billions of global verifications: the diversity and depth required for broad regional coverage and resilience against a wide range of fraud tactics.

4.1B+ IDENTITY CHECKS	**400M+** UNIQUE IDENTITIES	**4,600+** DOCUMENT TYPES	**200+** COUNTRIES COVERED
700+ ENTERPRISE CLIENTS	**20+** INDUSTRIES SERVED	**670+** IDENTITY DB CONNECTIONS	**15+** GOV BIOMETRIC SOURCES

Labeling pipelines

Human and automated review across millions of records. 200+ human labelers create training data and measure AI performance for each customer.

Synthetic data

120+ synthetic-data generation tools produce tampered documents, presentation attacks, and deepfakes to train models on rare threats.

Fraud Lab

A red-team environment that simulates and replays real-world attack scenarios, continuously stress-testing models against new fraud tactics.

03 · TRUST GRAPH

Fraud intelligence across organizations

A privacy-preserving graph that links identity signals across hundreds of millions of otherwise separate records held by governments and enterprises. By connecting these signals safely, it surfaces serial fraudsters and organized rings that no single company could see, and every check increases the density and diversity of training data, making the models stronger over time.

- Sub-20ms search on hundreds of millions of vectors
- Zone-level resilience and replication
- Encrypted, retention-controlled, fully auditable

04 · GOVERNANCE

Responsible AI, by design

Data practices	Purpose-limited, minimized, encrypted data with regional compliance options.
Dataset quality	Curated, balanced datasets with pseudonymization and continuous QA.
Fairness & bias	Bias testing across demographics with remediation and monitoring.
Monitoring & feedback	Real-time dashboards, drift detection, fraud-focused alerts.
Incident & continuity	24/7 monitoring, DR readiness, fast response to new fraud vectors.
Access & security	Role-based controls, secure SDLC, HSM key management, full audit logging.
Model development	Reproducible pipelines, versioned training, performance-driven tuning.
Deployment controls	Staged rollouts, canary checks, kill-switches, secure deploys.
Retention & deletion	Configurable retention, verified deletion, GDPR/CCPA-aligned.
Compliance	SOC 2 and ISO certified; GDPR, CCPA, LGPD compliant; public Trust Center.

Deep learning models and third-party validation

AREA	WHAT IT DOES	THIRD-PARTY VALIDATION
Face intelligence Face Detector · Recognition 1:1 · Recognition 1:N · FaceDB vector engine	End-to-end face perception: detects faces, builds robust embeddings, and matches identities at scale, improving continuously through calibration and hard-case mining.	NIST #1 for facial recognition · 1:1 and 1:N NIST certified · FIDO Face certified · DHS RIVTD benchmarks met
Liveness Face Liveness · Document Liveness	Multi-modal defenses that tell real users and physical documents apart from presentation attacks and replays, using spatial, temporal, and device-aware signals with continual hard-negative training.	First passive liveness technology certified in the market · iBeta ISO/IEC 30107-3 PAD Level 2
Deepfake defense Deepfakes · Gen-AI Documents	Multi-modal models that detect and block AI-generated fraud: deepfakes, face swaps, document injections, and synthetic identities.	#2 in the ICCV 2025 DeepID Challenge · #1 in deepfake attack detection, Hochschule Darmstadt
Age assurance Age Estimation	Policy-ready age estimation with calibrated predictions, uncertainty bounds, and fairness constraints; edge cases route to secondary verification.	NIST top 3 for lowest average error · NIST fastest response time · ACCS accredited under PAS 1296
Document intelligence Type Classification · Alteration & Tampering · Text Readability · Cropping · Barcode Validation	Classifies document type, extracts and validates OCR, MRZ, and barcodes, and detects tampering, fusing signals into an authenticity score that adapts with active learning.	Evaluated on large global identity-document datasets across 200+ regions
Fraud & risk Risk AI Agent · Trust Graph · Evasion Fraud	A real-time orchestration layer that combines model outputs with fraud-network intelligence and AI risk agents to score and route risk decisions.	250+ signals fused per identity check · Enterprise-grade accuracy with built-in interpretability
Device & behavior Behavioral Model · Device Signal Model	Analyzes device environments and interaction signals: blocks injected or emulated environments and flags scripted, non-human patterns.	Detects emulators, virtual cameras, hardware spoofing, and scripted automation

See it live: request a demo at **incode.com/contact**, or explore the platform at incode.com/technology.