

# Insurance Fraud Detection Dataset

## IFDD v1.0 — Dataset Fact Sheet

**1,000,000**

Records

**127**

Fields

**12**

Domains

**97.1**

Cert Score

### Source Intelligence

Source Type	Pct	Primary Sources
OIG_LEIE_SOURCE	18.0%	OIG List of Excluded Individuals/Entities
CMS_PI_SOURCE	20.0%	CMS Program Integrity, ZPIC, MAC reviews
DOJ_FCA_SOURCE	15.0%	DOJ False Claims Act enforcement actions
CMS_RAC_SOURCE	12.0%	CMS Recovery Audit Contractor findings
NHCAA_SOURCE	8.0%	National Health Care Anti-Fraud Association
STATE_FRAUD_SOURCE	10.0%	State Insurance Fraud Bureaus
DERIVED_INTELLIGENCE	17.0%	Fraud models, risk scores, predictive analytics