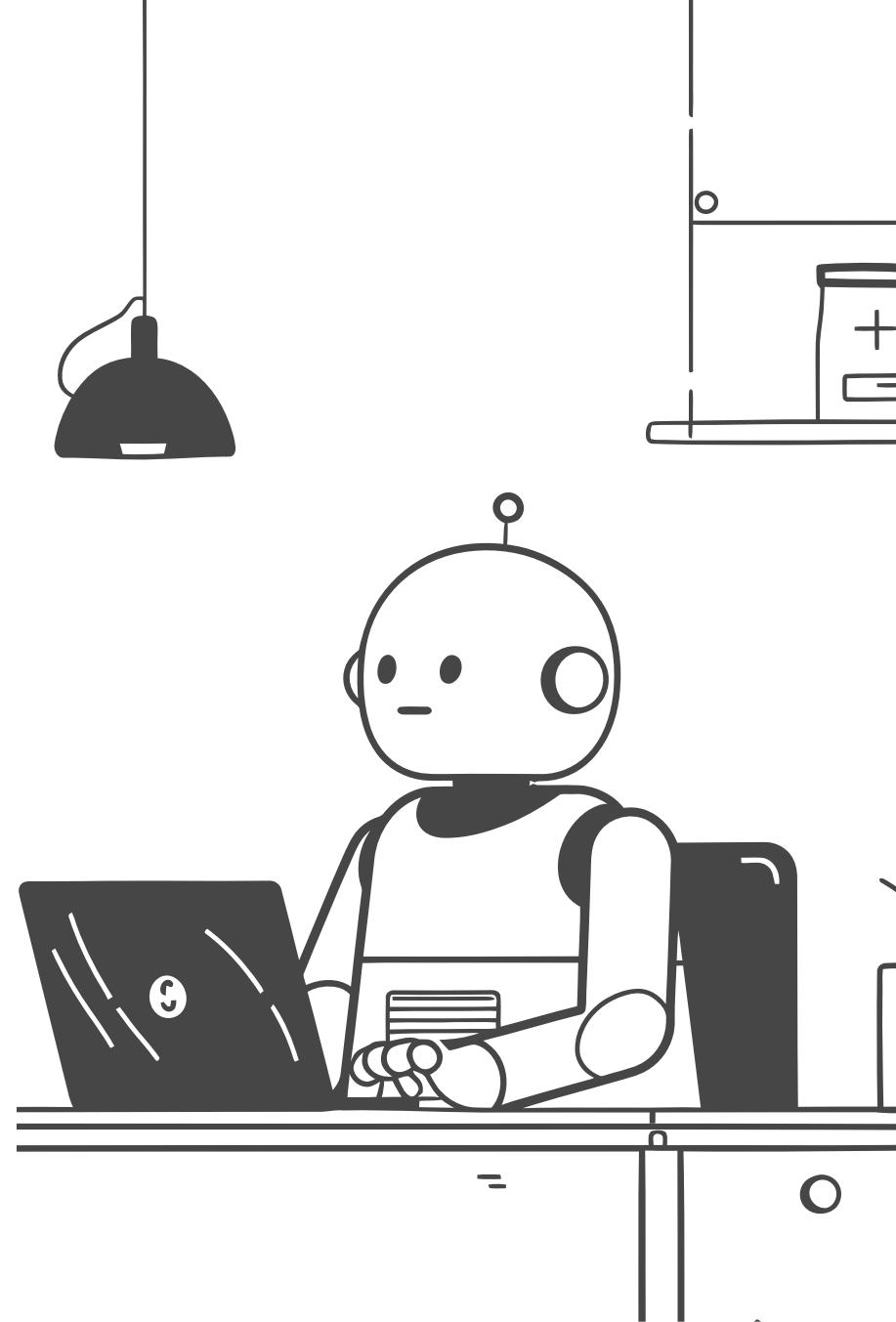


# Market Anomaly Detection

Andrea Clark | December 9, 2025

Presenting an agentic system for market anomaly detection built on structured WRDS data and natural language reasoning.



# Demo



YouTube

Market Anomaly Detection Demo



# Multi-Source Market Anomaly Detection Engine

An agentic system powered by IBM Agentics that detects market anomalies by integrating structured WRDS data, analyst expectations, and LLM reasoning to deliver actionable insights.



## WRDS Data Ingestion

CRSP daily prices, Compustat fundamentals, IBES analyst expectations, and CIQ KeyDev corporate events.



## Text2SQL Interface

Natural-language financial querying via schema-aware SQL generation—no manual queries required.



## Anomaly Detection

Independent signal channels unified into a z-score-based anomaly index with transparent components.

# Data Architecture – WRDS Integration Pipeline

## CRSP Market Data

Returns, volatility, and liquidity signals for comprehensive market coverage.

## Compustat Fundamentals

Revenues, margins, and leverage ratios for financial health assessment.

## IBES Analyst Data

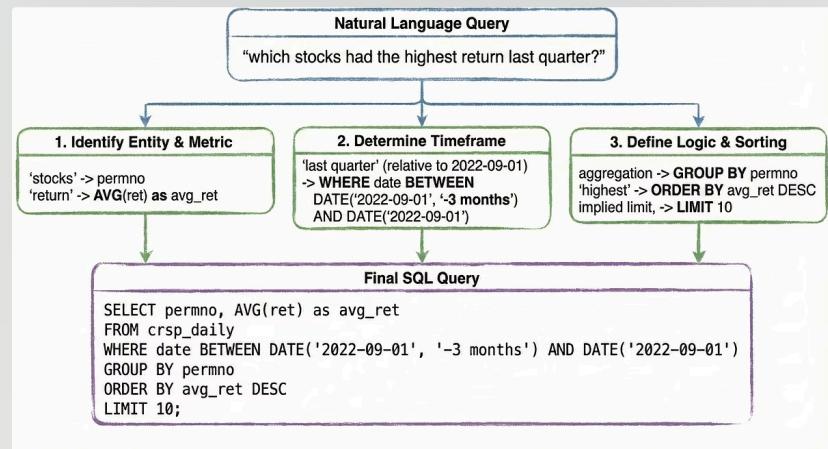
EPS forecasts, revisions, and surprises tracking market expectations.

## CIQ KeyDev Events

Executive changes, M&A announcements, and deal cancellations.

- All ingestors run through quality-controlled connectors and save to local Parquet storage, feeding feature generation and anomaly scoring.

# Text2SQL – Natural-Language Querying



01

## LLM SQL Generator

Schema-constrained, safe queries generated automatically.

02

## Schema Checker

Validates against CRSP, Compustat, and IBES columns.

03

## Validator

Column whitelist with auto-date clamping using latest CRSP date.

# Anomaly Detection & Scoring Framework

Channel	Core Signals	Examples
CRSP Statistical	Return shocks, volatility jumps, liquidity drops	Flash crashes, volume spikes
Compustat Fundamentals	Margin shifts, leverage surges	Profit warnings vs peers
IBES Sentiment	Forecast revisions, surprise dispersion	Analyst downgrades
CIQ / BoardEx	Executive changes & M&A events	CEO turnover, withdrawn deals

## Z-score Interpretation

- $|z| \leq 0.5 \rightarrow$  normal
- $|z| \approx 1.5 \rightarrow$  moderate (~33%)
- $|z| \approx 2.0 \rightarrow$  elevated (~50%)
- $|z| \geq 3.0 \rightarrow$  strong (~100%)

## Aggregate Score

$$A_{final} = w_1 A_{crsp} + w_2 A_{fund} + w_3 A_{sent} + w_4 A_{events}$$

Dynamic weights reflect each channel's information content across securities.

# Future Directions



## Integrate More External Signals

- FRED macro indicators (VIX, rates, spreads)
- Yahoo Finance benchmarks
- Google Trends search-volume spikes
- Align news timestamps with CRSP/IBES events for causal consistency



## Bayesian Market Priors

- Build an adaptive Bayesian anomaly model
- Have system learn prior distribution over typical market behavior
- High volatility → weaker prior (more flexible)
- Stable markets → stronger prior (shrinkage towards historical norms)



## Feature Expansion

- Add sector-normalized features: peer deviation, sector-normalized anomalies
- Explore temporal models (RNNs/Transformers) for sequence-aware anomaly scoring