FORENSIC ALPHA

Forensic Alpha API Solution

> "Though the footnotes are not nearly as readable as the Chairman's letter, they are far more important. Many a financial disaster has been caused by ignoring what was sitting in plain sight in the footnotes to the financial statements."
>
> *Warren Buffett's letter to shareholders, 1985*

A single company publishes thousands of pages of financial information every year. Yet conventional data providers map less than 1% of that information. Using Large Language Models to analyse unstructured data, Forensic Alpha unlocks valuable insight by accessing the other 99%.

Confidential

Capturing Unstructured Data

A single company publishes thousands of pages of financial information every year, all of it prepared according to accounting standards that require a 'true and fair' representation of the business. However because this information is largely embedded in paragraphs of text or unstandardized schedules, the vast majority of it is not captured by conventional data providers and cannot be analyzed in a systematic way.

Forensic Alpha uses Large Language Models (LLMs) to analyze the schedules, disclosures and footnotes presented in the notes to the accounts. These disclosures provide key information about the health and risks of the company, and often contain early warning signs of trouble. This might include the use of factoring to enhance reported cash flow, the establishment of a new off-balance sheet arrangement or loans made to a related-party

Alpha-Driven

Headline numbers, like Revenue and EBITDA, are heavily scrutinized by the analyst community. Because of this companies will often 'manage' these results and attempt to present the company in the best possible light. Where there is bad news or problems brewing, this information is usually disclosed in the footnotes or hidden towards the back of the accounts.

Our approach is focused on downside risk, which receives far less attention from the market than upside potential. For this reason, we believe it represents a stronger source of untapped alpha generation.

A Systematic Approach

With a purely automated approach, Forensic Alpha can systematically extract and analyze data from hundreds of new filings each day. For US stocks, we pull new filings in real-time, updating our scores within 30 minutes of publication. For European stocks, we update our scores within 24 hours.

Having standardized the data, we can make meaningful comparisons across quarters and across peers, allowing us to easily identify patterns and anomalies in the data as they arise. Our panel of 27 accounting flags captures the key risks from the financial filings, scoring each flag from 0 to 5. These individual flags are then aggregated up into an overall “Risk Score” from 0 to 10.

Proven Results

Our approach is built upon a substantial body of work performed by practitioners and academics in the field of forensic accounting and fraud detection. Academics such as Professor Messod Beneish pioneered the use of ratio analysis in fraud detection in the 1990s. Since then others have built on his work, incorporating a wider range of accounting and governance metrics and utilising machine learning approaches to uncover new patterns. Meanwhile, short-sellers such as Jim Chanos and Muddy Waters have long used forensic analysis to identify companies that may be engaged in accounting fraud.

While building on this body of knowledge to create our framework, we have also backtested our approach. Our backtest is based on a 12 year history for the Russell 1000 and shows cumulative performance of [X]% on a market-neutral long-short portfolio. Please see our separate white paper for further details.

API Overview

We have a range of APIs based on the use case, as detailed below. Note, our backtest is based on the “AccountingFlags API”

AccountingFlags API

The AccountingFlags API provides accounting scores for each stock under Forensic Alpha’s coverage. The “Overall Score” has a range of 0(Low Risk) to 10 (Very High Risk) and is constructed from our panel of 27 accounting flags. The methodology is consistent with the backtest

Use it to:

- Incorporate the signal from the Overall Score into your model
- Get the breakdown of the “Overall Score” into its constituent flags

RedFlags API

The RedFlags API provides accounting, governance and earnings quality scores for each stock under Forensic Alpha’s coverage. The “Overall Score” has a range of 0(Low Risk) to 10(Very High Risk) and is constructed from our panel of 27 accounting flags, 7 governance flags and 4 earnings quality flags.

Use it to:

- Incorporate the signal from the Overall Score into your model
- Get the breakdown of the “Overall Score” into its constituent flag

RedFlagsWording API

The RedFlagsWording API provides direct access to the wording used in our live reports on our portal.

Use it to

- Programmatically pull the latest wording for any specified stock and any specified module

Report API

The Report API returns a URL to the forensic report (as published on our portal) for a specified stock

Use it to:

- Obtain and easily distribute the most recent pdf report, as published on the Forensic Alpha portal.

RawData API

The RawData API provides access to the raw data points that are collected from the filings and analyzed by our system. Each module contains a number of raw data points that drive the score and can be accessed via this API.

Use it to

- Perform your own analysis of the raw data from the filings that feeds into our scores.

FilingExtract API

The FilingExtract API enables you to retrieve the text of specified sections from a specified filing. The section is rendered in markdown format, and where the information is in a non-English language, it will be translated into English. Note currently the FilingExtract API is only valid for stocks outside the US

Use it to:

- Programmatically extract sections of interest for further analysis
- Create a reliable feed for a LLM pipeline

Coverage API

The Coverage API lists every company currently covered by Forensic Alpha

Use it to:

- Identify stocks under coverage within a given region (US, Europe, Asia-Pacific)
- Check the coverage status of a specified stock
- Map between different identifiers (isin, cid, ticker)

Dataset description:

Accounting red flags, sourced from structured and unstructured data in public filings

Asset class: Public equities
Data sources:
SEC Edgar (US Equities), annual and interim reports (European Equities)
Regional coverage:
US: Coverage of 2,350 companies over $1bn market cap
Europe: Coverage of 1,450 companies over EUR 500m market cap
Asia: Coverage of 120 companies, in Hong Kong, Singapore and Australia. Set to expand to other exhanges in H2 25"
Sector coverage:
All sectors ex Banks, Insurance
Data history:
Historical data for Russell 1000 index since 1/1/2013 (12.5 years)

Dataset: Delivery and technology

File format: JSON, CSV

Available identifiers: ISIN, CUSIP, Bloomberg Ticker, CID

Frequency of updates: Real-time in US, Daily outside US

Supporting Documents (Available on Request)

- Full dataset guide
- Backtest whitepaper
- Technical API documentation
- Coverage list

Contact details

Name: Thomas Beevers

Email: tom@forensicalpha.com

Tel Number: +44 (0) 779 669 2275