



How Entity Recognition Enhances Five Finance Workflows



Introduction

In financial markets, decisions hinge on timely, precise understanding of entities - companies, people, products, institutions. MarketPsych's Text Tagging Engine uses advanced Named Entity Recognition (NER) to extract point-in-time entities from financial text such as news, research, and filings. This enables cost-effective and accurate search, alpha-generation, and generative AI. Five successful NER use cases from our clients are described below.

1. Accurate Entity Tagging (Financial Media)

yahoo!finance

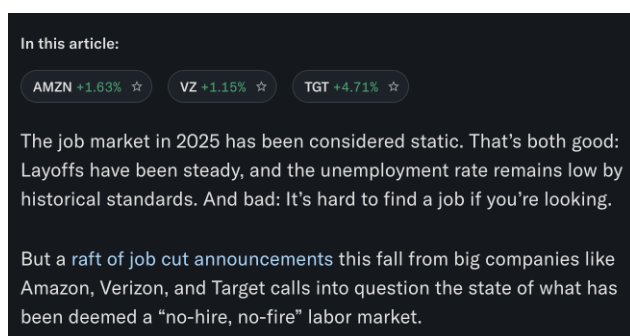


Figure: Note the link between news entity references and real-time price change data, made possible with the NER module of the Text Tagging Engine.

Challenge:

Yahoo! Finance tags tens of thousands of companies (both public and private), commodities, currencies, cryptocurrencies and link them to market prices. But they have a global news collection in 28 languages and must deal with the challenge of changing and ambiguous company names and identifiers - IPOs, mergers, splits, cashtags, tickers, and non-English names. They evaluated multiple NLP products before choosing MarketPsych's NLP as best-in-class.

Solution:

Their evaluation environment tested each model on:

- Entity recall and precision for all types of tradable securities (ETFs, indexes, commodities, etc...)
- Handling of ambiguous names (e.g., "Voyager," "Unity," "Amazon") in financial news
- Speed of incorporation of new names (following IPOs, corporate spin-offs, mergers)
- Non-English language identification of corporate names
- Robustness to noisy text and document structures

Outcome:

Yahoo! Finance found that MarketPsych's Text Tagging Engine provided the highest-quality NER of any product tested. Solution deployed on both Yahoo! Finance web pages and iPhone's Stocks App.

- Articles were tagged more consistently and correctly.

- Fewer mis-tagged stories appearing under the wrong ticker page.
- All types of assets (ETFs, Indexes, Cryptocurrencies, etc...) were correctly tagged.
- As a result, there was an improved user experience and higher engagement on news pages.

2. Research & Real-Time Alerts (Large Global Bank)

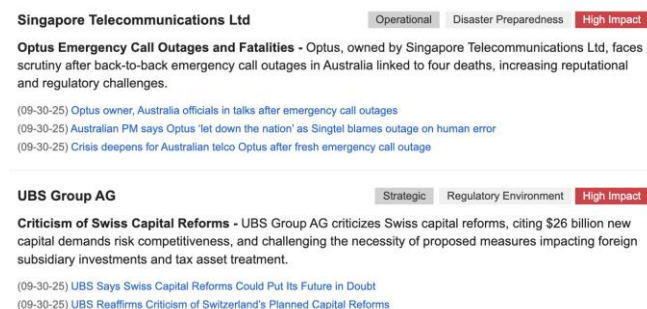


Figure: Examples of "High Impact" notifications from a real-time NLP-based alerting pipeline.

Challenge:

A Tier-1 global bank needed a way to track risk-relevant changes across thousands of internal and external news, transcripts, and research reports. Analysts were manually searching documents to identify key entities - executives, regulatory bodies, facilities, and risk topics - but they relied on poor entity and topic recognition. Traditional keyword search frequently missed documents where synonyms or historical company names were used.

Solution:

The bank integrated the Text Tagging Engine's **entity recognition** and **topic-tagging** module through its document ingestion pipeline. The engine tags each article and sentence with point-in-time entities, topics, events, and sentiment.

Outcome:

- They can now generate **automatic alerts** when documents contain material mentions tied to high-risk topics (e.g., litigation, credit downgrades, management changes) across sectors and geographies.
- There was a **90% reduction** in time to extract key topics and themes
- They receive ESG and credit-risk references fast, direct to their inboxes.
- Time-normalized references are utilized to assess urgency (e.g., "next quarter" → specific date range)

3. Alpha Generation from Text (Quant Hedge Fund)

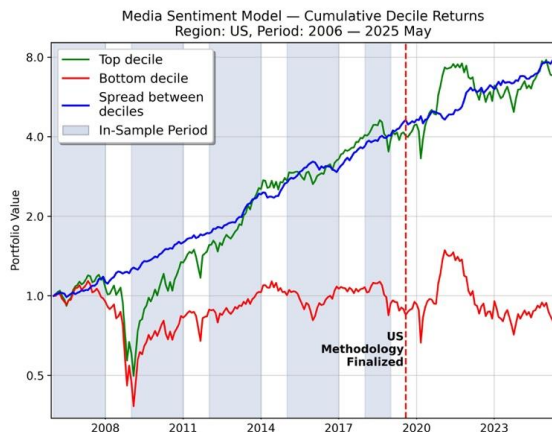


Figure: The equity curve of a predictive model designed using NLP, which is dependent upon accurate point-in-time entity recognition.

Challenge:

A multi-strategy quant hedge fund sought untapped alpha signals from corporate press releases, earnings announcements, and specialized industry publications. Their existing NLP stack captured sentiment, but its entity detection was difficult and expensive to keep up-to-date, and it frequently failed to resolve historical or ambiguous company names.

Solution:

The firm integrated the Text Tagging Engine into its daily data processing workflow to:

- Identify point-in-time entities at both sentence and document level, including upcoming IPOs, and link them directly to security identifiers.
- Tag granular financial events such as EARNINGS-INCREASE, GUIDANCE-CUT, PRODUCT-LAUNCH, CREDIT-RATING-DOWNGRADE.
- Combine entity-linked event tags with sentiment outputs to build event-driven return prediction models.

Result:

- The Text Tagging Engine enabled faster deployment of thematic (e.g., robotics, clean energy, AI hardware) and risk-related strategies.
- They improved the precision of their backtests by removing survivorship bias.
- They produce orthogonal alpha factors that are additive to their traditional quant factors.

4. Credit Risk Monitoring (Financial Regulators)



Figure: A view of JSON output from a headline describing allegations of sexual exploitation against Meta.

Challenge:

A central bank monitors 25,000+ financial institutions globally. They needed automated detection of references to risk events: credit stress, fraud accusations, fines, layoffs, sanctions, insider trading, and executive departures.

Solution:

By connecting inbound news feeds and internal documents to the Text Tagging Engine, the bank automatically extracts entities and associated risk topics and sentiments:

- Risk-related event tags
- Negative sentiment indicators
- Geopolitical entity (GPE) relationships
- Facility-level mentions

Outcome:

- Real-time dashboards highlight emerging risk clusters.
- Better early warning detection for clients with increasing negative media pressure.
- Credibly auditable trail of entity-linked evidence.

5. Generative AI Summaries (Brokerage)

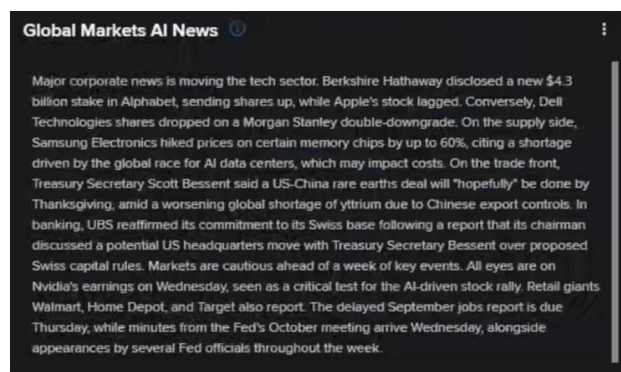


Figure: A summary of daily economic and market news covering the AMERS region. All regions and national markets are covered

Challenge:

A major brokerage needs an automated daily briefing that summarizes relevant news for each asset in a client's portfolio. Traditional LLM summarization tools are too generic and often included irrelevant items.

Solution:

The Text Tagging Engine pre-processes all incoming documents, identifying entity relevance at sentence and article level. These structured tags feed a lightweight LLM, producing precise, client-specific summaries.

Outcome:

- Personalized portfolio summaries without hallucinations.
- Reduced LLM workload and lower compute cost by ~60%.
- Investors receive highly relevant, client-specific intelligence.

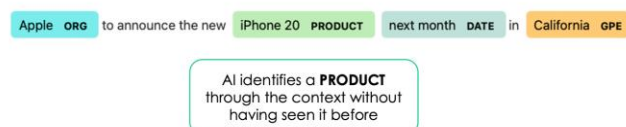


Figure: A view of the granularity achieved with LLM-driven entity recognition.

Summary

The NER Component of MarketPsych's Text Tagging Engine is the key to its reliability and cost-savings, and it comes with unique benefits:

- **Custom Financial NER:** Unlike generic NER models, our engine is trained on a **rich, finance-specific taxonomy** and identifies millions of entities of 20 classifications (e.g., companies, products, currencies, central banks, institutions, etc.).
- **Point-in-Time Awareness:** Our system **supports temporal entity resolution**, correctly linking entities as they existed at a specific time (e.g., historical company names, merged entities), preventing survivorship bias and enhancing alpha generation.
- **Generative AI & Retrieval-Augmented Workflows:** The system can feed structured entities into **RAG (retrieval-augmented generation)** pipelines, enabling generative AI with citation, grounded in reliable source documents — reducing hallucinations and improving answer precision.
- **Scalable, Real-Time Processing:** Our GPU pipeline is built to ingest and process **millions of documents daily**, with low latency and robust performance to article surges.
- **Granular Tagging:** Fine-tuned and specialized LLMs tag date references, dependency relationships, themes, sentiments, and more which are then assigned to each associated entity.

Please contact info@marketpsychdata.com or use our Contact form for more information on how we can bolster your NLP efforts.