



Trending Themes

Emerging media themes, structured as quantitative data



Overview

Trending Themes detects statistically significant surges in global media coverage as they develop, and maps each theme to sentiment, relevance, and urgency scores across global entities. Delivered daily via API.

Emerging media themes are structured as **quantitative data**: every theme is tagged with a narrative summary, category, and a complete set of entity-level scores so quantitative and discretionary workflows can act on developing narratives within hours of formation.

145K+

Mapped Entities

28

Languages Monitored

1,000+

Curated Media Sources

2005

Historical Coverage From

How It Works

A multi-stage pipeline combining classical NLP with large language models

The pipeline surfaces statistically anomalous narrative clusters and maps them to entities. The encoder-based NLP layer handles statistical detection; a generative AI layer labels and summarizes the output for human readability.

01 Text Ingestion

Millions of documents published across 1,000+ curated news and social media sources are processed daily through an NLP engine to uncover key phrases, entities, and sentiment scores.

02 Surge Detection

Three statistical methods identify key phrases with anomalous chatter volume. A phrase is flagged only when at least two methods agree, filtering out low-signal chatter.

03 LLM Theme Labeling

A large language model reviews the sentences surrounding the surging phrases and assigns a theme name, a narrative summary, and relevance and urgency scores.

04 Entity Mapping

Each theme is linked to the companies, countries, currencies, commodities, and other entities mentioned in its sentences. Every entity receives an individual sentiment score scoped to that theme, with a buzz count.

05 Delivery via API

A structured output in JSON or CSV format is published daily via RESTful API with historical access available from 2005.

Figure: The five-stage Trending Themes pipeline, from raw ingestion through structured API delivery.

Theme	US imposed tariffs
Category	Economics & Macro
Top Entity	AAPL
Entity Sentiment	-0.96
Theme Buzz	139 sentences
Urgency	0.90

Figure: Example daily record joining a theme to its top-ranked entity.

Real-World Examples

Themes detected as they emerged

These examples were detectable in the Trending Themes feed as the media clusters formed around each event, often hours or days before they appeared in mainstream coverage summaries.

NEGATIVE CONTAGION

Silicon Valley Bank Collapse, March 2023

As SVB coverage intensified, the feed detected a cluster of sentences co-mentioning SVB with counterparties holding deposits. Linked entities included potential acquirers and firms with similar risk profiles.

[SIVB](#) · [First Republic](#) · [Signature Bank](#)

MARKET DISRUPTION

DeepSeek GPU Efficiency Narrative, Jan 2025

DeepSeek coverage triggered a statistically significant surge in negative-sentiment sentences associated with S&P 500 semiconductor names. The theme was detected as the media cluster formed, allowing quant models to respond on day one.

[NVDA](#) · [S&P 500](#) · [SMCI](#) · [AMD](#)

CROSS-ASSET SPREAD

US Tariff Escalation, 2025

As tariff references in earnings calls reached multi-year highs, Trending Themes mapped the sentiment to individual equities, separating companies with positive associations (domestic manufacturers) from those with negative associations (importers, multinationals).

[XOM](#) · [NKE](#) · [AAPL](#) · [USD/CNY](#)

POSITIVE ASSOCIATION

GameStop Short Squeeze, 2021

The GME squeeze produced a secondary theme: media co-mentioning of Evotec SE alongside GameStop created a statistically anomalous positive association. The entity mapping layer captured this cross-asset signal as the cluster emerged.

[GME](#) · [Evotec SE](#) · [AMC](#) · [Reddit](#)

Data Format

Two endpoints. Complete coverage.

Every daily delivery consists of two main structured datasets: **Themes**, one row per theme; **Entities**, one row per asset per theme. Below, one finds a sample of the available fields in each endpoint.

GET /themes

Field	Description	Example
Date	Timestamp of theme.	2026-03-19T11:00
Theme ID	Persistent alphanumeric ID.	CMBKXGE26SZQA
Theme	LLM-generated short name.	US imposed tariffs
Category	Overarching category.	Economics & Macro
Summary	LLM narrative summary.	The shortage in GPU...
Emerging Days	Days since trending.	2
Buzz	Sentence count per day.	139
Sentiment	Avg sentiment, -1 to +1.	0.32
Relevance	Market impact score.	0.80
Urgency	Attention demand score.	0.90

GET /entities

Field	Description	Example
Entity	Identifier (name/ticker/ID).	AAPL
Theme ID	Persistent alphanumeric ID.	CMBKXGE26SZQA
Buzz	Sentence count in theme.	17
Sentiment	Avg sentiment in theme.	-0.96
Relevance	Entity weight in theme.	0.80

Use Cases

Applications across the investment workflow

Trending Themes data fits quantitative, discretionary, risk, and product-construction workflows. Below are common applications.

Quantitative Trading

- Generate alpha signals from statistically anomalous narrative clusters before they reach mainstream summaries.
- Filter trade ideas by relevance and urgency scores to focus on high-conviction themes.
- Build cross-asset spread strategies from theme-level entity sentiment.

Discretionary Research

- Surface emerging narratives across coverage universes in a single daily digest.
- Track theme persistence with the Emerging Days field to separate noise from sustained narratives.
- Map theme exposure across portfolios using entity-level Buzz and Sentiment.

Risk Management

- Detect early-stage contagion signals such as the SVB cluster days before broader market repricing.
- Monitor portfolio-level theme exposure with daily entity-mapped sentiment scores.
- Backtest theme-driven drawdown signals against historical coverage from 2005.

Thematic ETF & Product Construction

- Identify emerging themes early to inform new product launches and index construction.
- Validate theme persistence and breadth using buzz and entity-mapping data.

Get Started

Request a data sample or trial access

Contact us to set up a proof of concept for your asset universe. Sample data files and API access are available on request.

marketpsych.com/contact-us
sales@marketpsychdata.com

Summary

Trending Themes combines statistical surge detection with large language model labeling to deliver emerging media narratives as structured, quantitative data. With entity-level sentiment, relevance and urgency scores, the feed is built for quant trading, discretionary research, risk management, and thematic product construction. Daily delivery via RESTful API with historical access from 2005.