



Global

News Analytics Dataset for Stocks



Swiss Re

mazars



CREDIT SUISSE



example clients

About the Dataset

This dataset offers end-of-day data with different News Analytics Data points, derived from global news analysis, such as Sentiment, Risk and ESG.

For deeper Analysis, it also contains the exact amount of underlying events, for each ESG, Risk or general events.

About Us

We are YUKKA Lab, Berlin-based Al pioneers in news analysis. Our technology processes over 1,000,000 daily news items from 200,000+ sources, transforming them into actionable metrics for risk management, ESG, investment, and sales.

Key Facts and Figures

- Various End-of-day News-analytics data per company
- Covers over 200,000 private and public companies
- Updated daily at midnight GMT
- Accessible via API for seamless integration
- SaaS solution available for in-depth analysis
- Quant data starts in 2000, with full release starting in 2020

Your Benefits



Real-Time News Data

- Provides timely company-related data based on global News
- · Enables proactive management and risk mitigation



API Access

- Seamless integration with existing systems and workflows
- Flexible and scalable access to news data
- Enhances operational efficiency through automation

Quantitative Modelling

- · Enables creation of investment strategies using ESG, Risk and News data
- · Facilitates detailed quantitative analysis and modelling



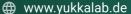
Alternative Delivery Methods Newsletter

- · Alerting Service
- PDF reports



+49-30-610 818 69-0





Data Schema

Name	Field	Description Data Type		Scale
Date	date	Datestamps based on GMT Timezone Timestamp		-
ISIN	isin	International Securities Identification String Number		-
Internal Entity ID/Name	entity	Unique ID used internally by YUKKA String		-
Positive	positive	Total numbers of positive news Instances Integer		0+
Negative	negative	Total numbers of negative news Instances	Integer	0+
Sentiment	sentiment	Total daily news sentiment, based on the number of positive and negative Instances	Float	0 (Negative) - 1 (Positive)
Volume	volume	Total numbers of news articles on a given date mentioning the company	Integer	0+
ESG Score	esg_score	A real-time score for ESG that is based on mid-term trends (weekly to monthly)		
E Score	e_score	A real-time score for Environment that is based on short-term trends (daily to weekly)	Float	0 (Bad) - 10 (Good)
S Score	s_score	A real-time score for Social that is based on mid-term trends (weekly to monthly)	Float	0 (Bad) - 10 (Good)
G Score	g_score	A real-time score for Governance that is Float based on Mid-term trends (weekly to monthly)		0 (Bad) - 10 (Good)
ESG Risk Score	esg_risk_score	A real-time score for ESG Risk that is based Float on short-term trends (daily to weekly)		0 (Safe) - 10 (Risky)
Immediate Risk Score	immediate_risk_s core	A real-time score for Risk that is based on short-term trends (daily to weekly)	Float	0 (Safe) - 10 (Risky)
General Risk Score	general_risk_sc ore	A real-time score for Risk that is based on Float mid-term trends (weekly to monthly)		0 (Safe) - 10 (Risky)
Credit Risk Score	credit_risk_score	A real-time score for Credit Risk that is based on short-term trends (daily to weekly)	Float	0 (Safe) - 10 (Risky)
Event Count	event_counts	one event_counts field for each of the relevant events	Integer	0+



Example data

date	entity	ISIN	esg_score	general_risk_score	 Lawsuit	Industrial Accident
2024-06-01	company:abc	DE123 456	8.827	8.861	 1	3
2024-06-01	company:def	DE78 910	2.780	9.816	 42	

Description

The columns date and entity function as multi-index. Score values in a row mean that the respective score for that company at the end of that day is X. Event Values like Lawsuit signals that on that day, this company was detected N times as Participant of Lawsuit. The sample shows only a small subset of all available Events for brevity.

Score Calculation

The Scores are calculated, based on the detected events and news volume of a company. Each score considers a subset of our 160+ events with a different weight and calculates this score based on the amount of risk events in the last 100 days. The Score is normalized and standardized to be uniformly distributed across all companies.

Sentiment Calculation

The Sentiment S is calculated as S=(P+1)/(P+N+2), where P is the amount of positive and N is the amount of negative mentions.

