

INTEROS // WHITE PAPER

Interos: A Technical Overview

Version 2.0, November 2019

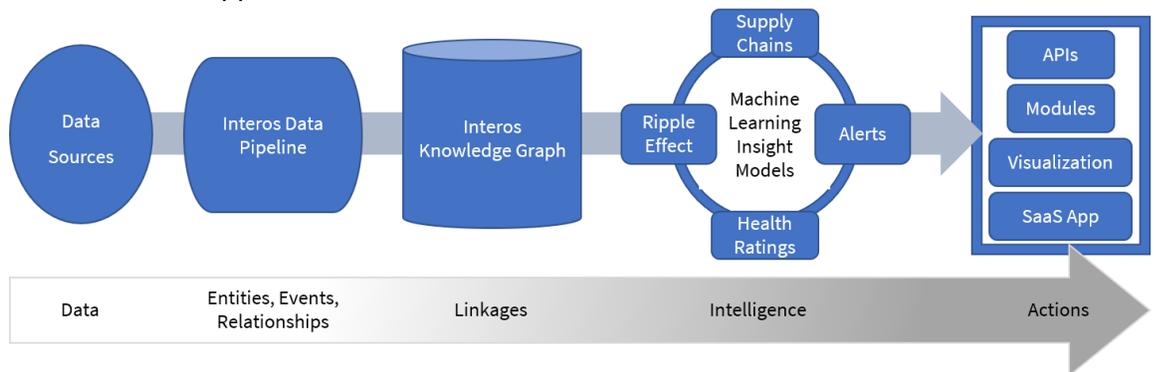
WWW.INTEROS.AI

Overview of the Interos Platform

The goal of the Interos Platform is to help our customers improve the resiliency, compliance, and value of their supply chain and business partnerships or, when combined, their ecosystem.

To deliver on this goal, the core interaction of Interos facilitates the discovery and assessment of risk in supply chains and business relationships and the impact (or ripple effect) of those risks as they propagate throughout our customers' ecosystems. Supporting this core interaction requires the continuous extraction of millions of business entities, business relationships, products, timeline events and derived insights from petabytes of dynamic data.

In addition, to ensure our core interaction is responsive to the dynamic nature of the macro-business environment, we utilize a wide range of machine learning algorithms to uncover and infer relationships and insights. To bring the core interaction to life, the Platform merges our proprietary knowledge graph, data pipeline, machine learning and statistical models, APIs, insight-centric modules, data visualization and our SaaS-based application.



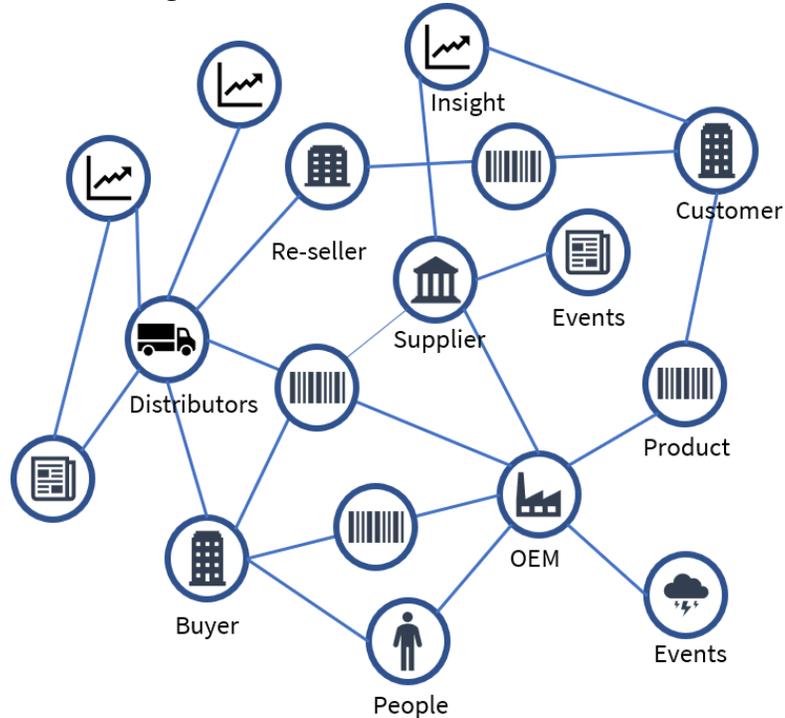
Interos Knowledge Graph

The brain of the Platform is the Interos Knowledge Graph which captures:

- Business, product and people entities,
- Business relationships such as “supplier of” and “customer of”,
- Timeline events such as a company’s revenue and natural events,
- Attributes such as role and location.

Today, the graph encompasses over 5m business relationships, 1.2m supplier relationships, 50m businesses and 250 million monthly events. These entities, events, relationships and the linkages between them form the ontology of a real-world business ecosystem that defines how businesses interact, interconnect and which businesses an event impacts. Maintaining this real-world ecosystem requires that our Knowledge Graph continuously captures the daily activities that impact business ecosystems. As a result, Interos architected its Platform to continuously add and link new entities, attributes and relationships to the Interos Knowledge Graph and to capture changes to existing entities, relationships and links in the graph.

The Interos Knowledge Graph



Linkages and Relationships

While some entity relationships are inferred by the Interos Knowledge Graph, the majority of the relationships are explicit and discovered by a large suite of machine learning and natural language processing algorithms. The graph links these individual entities and relationships together to form the graph's comprehensive ecosystem. For example, the Interos graph will discover and link the multi-tier relationship that Acme Corp's supplier for circuit boards is Zeta Corp and Zeta Corp's suppliers' components are from Beta Corp and Alpha Corp, and then re-link when Acme Corp switches to a new supplier.

Furthermore, the graph continuously discovers and links attributes and events that impact an entity and relationships. For example, the graph will detect that one of Acme Corp's tier 3 suppliers is impacted by a weather event, prompting Acme Corp to switch to a backup supplier. The graph reflects this change and illustrates that their new supplier has a tier 2 supplier in a sensitive country.

The Graph's Data Pipeline and Accuracy

Maintaining the dynamic nature of the Interos Knowledge Graph requires the continuous extraction, transformation and loading of terabytes of externally discovered unstructured, semi-structured, and structured data. The types of data we ingest include:

- Time-Series
- Economic
- Transactional
- Current Events
- Social
- Business Generated
- Financial
- Weather and Natural Disaster
- Social
- Regulatory
- Cyber
- Geographic

To ensure that this data, as well as the extracted entities, relationships, attributes and events exceed a minimum threshold of relevancy and confidence, Interos utilizes a multifaceted data processing pipeline. This pipeline starts with the inclusion of a new data source and ends with the extraction, transformation and insertion of new entities and relationships. The pipeline interrogates and scores the confidence level and accuracy of the data prior to inserting an extracted entity, relationship or event into the graph.

Before ingesting a new data source, an Interos analyst verifies and scores the quality of the source based on the reputation of the publisher, the origins of data sourced and other consumers of the source. Once a data source is validated, scored and included in our pipeline of over 85,000 sources, Interos utilizes a farm of web crawlers, API connections and RSS feeds to extract the data, and applies an algorithm (similar to page-ranking) to continuously score and monitor the source's quality. These quality scores are used in the downstream algorithms to aid in balancing the trade-off between accuracy, precision and coverage.

Maintaining High Quality Data

To solve the challenge of extracting entities, relationships and attributes from the 200 plus articles that we ingest every 20 seconds, Interos' data pipeline applies machine learning techniques and over 25 proprietary algorithms to filter spam, cleanse and standardize data, de-dupe, and disambiguate extracted entities, relationships, events and attributes. In addition, all entities, relationships, events and attribute extraction models have confidence scores that are computed by a machine learning model. When a result's score falls below a minimum threshold an analyst is signaled to curate and verify or correct that result. Depending on the type of model, we measure confidence thresholds using F1 scores, mean average error, mean squared error, accuracy, precision and recall. The following types of model and their thresholds are:

- **Entity Extraction Models** - Detects, disambiguates and extracts ecosystem entities and has a minimum accuracy threshold of 90%.
- **Entity Linking Models** - Links ecosystem entities' in-text mentions, discovering attributes and events and has an F1 score (a measure of recall and precision) threshold of 85%.
- **Relationship Models** - Extract relationships between companies, products and people and enforces a minimum precision of 95%.
- **Events Extraction Model** - Uses clustering and sentiment analysis to triangulate across multiple articles and data sources to validate that the extracted event is accurate and has a minimum precision threshold of 75%.

As a final step, the results and scores generated by the models, and the results corrected by analysts, are then used to continuously train and improve the models.

Knowledge and Core Interaction Derived from the Graph

Leveraging our insights and machine learning algorithms across the 25 plus terabytes of active data in our Knowledge Graph, Interos continuously derives customer-centric core interactions and insights that capture the dynamic nature of a customer's business ecosystems. Utilizing deep learning, machine learning and statistical learning algorithms, we derive insights on our customer's supply chains, business relationships, the resiliency of individual suppliers and partners, and the impact or "ripple effect" of an event on our customer.

Supply Chain

To derive a customer's supply chain, our Supply Chain Model queries our Knowledge Graph. The model collects the relevant individual relationships that live in our graph and links the relevant product entity and role attributes to each individual entity and relationship. This forms what we call a transaction. The Supply Chain Model then assembles and aggregates all the relevant transactions into a specific customer's multi-tier supply chain graph. Further, if our customer switches a supplier or if there is a change in our customer's supply chain, our Supply Chain Model automatically repaints their supply chain graph to reflect that change. To ensure the accuracy of our Supply Chain Model's output we require an 85% accuracy and 95% precision threshold.

Insights and Health Factors

Using a suite of ordinal classification and regression algorithms, we derive health ratings and insights that measure a company and the resiliency of its supply chain. To derive these health ratings, our Health Factor Model continuously analyzes and scores five health factors and over thirty variables extracted from the Interos Knowledge Graph for all companies in our customer's supply chain and business graph. These scores are then computed and recorded as discrete time-based events enabling us to provide trend and benchmark analysis on a customer-centric basis. To derive additional insights our models also inspect each variable and sub-variable to extract contextual insights that map to the health score. To establish confidence levels for scores, we measure the Root Mean Square Error (RMSE) for each variable and factor model and use that score in downstream models to calculate the overall health score or rating for a company.

The "Ripple Effect"

Central to our core interaction is what Interos calls the "ripple effect." The ripple effect is the probability of a business disruption caused by events as they propagate throughout the connections in a customer's supply chain and business relationships. Analyzing the ripple effect requires first that Interos derives a customer's supply chain to identify individual businesses and then continuously monitors those businesses to extract relevant events and changes in relationships and their potential impact. Running our Continuous Monitoring Model across our customer's supply chains, we extract, classify, score and link these events or risks to individual suppliers. Using our Alerting Model, we then notify our customers of an event that has a potential impact on an individual supplier. This model also generates the data needed for the downstream ripple effect model.

For example, since our Supply Chain Model has assembled Acme Corp's supply chain, our Continuously Monitoring Model now has the requisite data to identify and monitor the suppliers in Acme Corp's supply chain. In this case, the suppliers monitored include Zeta Corp, Beta Corp and Alpha Corp. So, when Beta Corp's (a tier 3 supplier) plant in Mexico is hit by an earthquake and Zeta Corp adds a new supplier in a sensitive country, our customer Acme Corp will receive an alert. To ensure our Continuous Monitoring Model is accurately classifying and scoring these types of events, we have minimum precision threshold of 75% with a target of 95% and an accuracy threshold of 90%. Like all of our models, when a result returns a confidence level below the threshold, the output is sent to an analyst to validate and, if necessary, correct.

Conclusion

Creating and maintaining a platform that captures the dynamic nature of a real-world, interconnected business ecosystem requires a robust knowledge graph and sophisticated data pipeline that continuously extracts businesses and relationships and links relevant events and attributes with a high degree of confidence.

However, that is only part of the equation. For customers to drive value from Interos and understand the impact of events and connections as they propagate across their ecosystem, we must also continuously run machine learning models on top of our graph to derive actionable insights. Ensuring and maintaining confidence in the accuracy and precision of our results is key to delivering actionable insights.