ISSN: 3067-4166
AMERICAN DATA SCIENCE JOURNAL FOR ADVANCED COMPUTATIONS
VOLUME: 03 ISSUE: 03
RECEIVED: JULY 08 REVISED: AUGUST 04 ACCEPTED: AUGUST 24 PUBLISHED: SEPTEMBER 12

A Unified Multi-Hazard Data Ingestion and Risk Inference Engine for Global Supply Network Stress Testing

Vikram Boga
Independent Researcher
bogavickram@gmail.com

Abstract

Supply chain risk prediction supports decisions on investment, protection, and recovery against disruptions. However, the prediction of such risks at a global scale remains underdeveloped. Artificial-intelligence-enhanced data engineering frameworks can contribute by defining relevant risks, specifying architectures and components for risk-analytics drivers, and identifying metrics for predictive performance. Informed by these contributions, AI-enhanced data-engineering frameworks offer a step towards meeting the prediction challenge.

A supply chain risk prediction framework elaborates general-purpose prediction pipelines using real-time data streams and multimodal data fusion for demand, supply, and logistics risks. Data pipelines for crisis risk signals incorporate geopolitical, economic, and weather-related predictors. Global pipelines remain sensitive to prediction quality for individual geographies. Hence, prediction performance metrics encompass accuracy, degree of calibration, area under the receiver operating characteristic and precision-recall curves, and risk-relevant decision-analytic metrics. Guidelines for robust, generalizable, and calibrated predictions cover stress testing, cross-domain evaluation, domain-adaptation concerns, and uncertainty quantification.

Keywords : Global supply chains; risk prediction; data engineering; artificial intelligence; machine learning.AI-Driven Supply Chain Analytics;Predictive Risk Modeling;Data Pipeline Automation;Machine Learning for Supply Chain Resilience;Real-Time Supply Chain Monitoring;Big Data Integration for Logistics;Intelligent Risk Forecasting;Distributed Data Engineering Architecture;Supply Chain Disruption Prediction;AI-Powered Decision Support Systems.

1. Introduction

Acute disruptions in global supply chains have underlined the lack of predictive capabilities as a key weakness. Data engineering frameworks for large-scale machine learning models must offer comprehensive pipelines down to the storage and processing of data features customized for prediction problems. These considerations lie at the core of three interrelated research problems addressing the prediction of supply chain risk signals driven by demand, supply, logistics, geopolitical events, or weather phenomena. The first problem focuses on three core supply chain risk indicators: demand shocks in consumer markets, supply shocks from important production countries, and freight rates representing logistics constraints. A second research angle assesses external risk sources beyond the immediate reach of supply chain actors—geopolitical positions, weather events, and macroeconomic factors whose influence on supply chains is documented but not yet operationalized in dedicated prediction models.

The third problem concerns event-driven risk signals that are manually documented but lack automated collection and presentation; facets such as military escalations, sanctions, pandemics, and natural disasters would therefore benefit from dedicated prediction models. An underlying hypothesis asserts that timely alerts for these external shocks could enhance decision-making in the supply chain context. Core domain knowledge is inherently required to define prediction methods, as the statement excludes topics such as sensor or image analyses for natural disasters that belong to the domain of the associated physical and computer sciences. Nevertheless, consider the domain of risk prediction, where assignable models are available for technological infrastructure events, new pandemic outbreaks, or major military tensions. Testing their predictive power,

ISSN: 3067-4166

AMERICAN DATA SCIENCE JOURNAL FOR ADVANCED COMPUTATIONS

VOLUME: 03 ISSUE: 03

RECEIVED: JULY 08 REVISED: AUGUST 04 ACCEPTED: AUGUST 24 PUBLISHED: SEPTEMBER 12

exploring the combination of different models, and investigating the development of structured prediction ensembles forms an open research agenda.

Event-driven risk signals remain a critical yet underdeveloped component of modern supply-chain risk management. Currently, many signals related to geopolitical tensions, sanctions, pandemics, and natural disasters are manually documented, resulting in delayed awareness and fragmented decision support. Automating the collection, analysis, and presentation of such signals could significantly improve the responsiveness of supply-chain stakeholders to external shocks. The central hypothesis is that early detection and timely alerts regarding disruptive events enable organizations to anticipate disruptions, adjust sourcing strategies, and mitigate operational risks more effectively. Developing such systems requires strong domain expertise to identify appropriate prediction methods while excluding areas outside the scope of supply-chain risk analytics, such as sensor-based or image-driven disaster detection that belong to physical and computer science domains. Within the field of risk prediction, however, several assignable models already exist for forecasting technological infrastructure failures, emerging pandemic outbreaks, and escalating military conflicts. Future research should therefore focus on systematically evaluating the predictive performance of these models, exploring how different forecasting approaches can be integrated, and designing structured ensemble frameworks that combine multiple predictive signals. Such ensembles could enhance robustness and accuracy, ultimately providing more reliable early-warning mechanisms for supply-chain decision makers.

ISSN: 3067-4166

AMERICAN DATA SCIENCE JOURNAL FOR ADVANCED COMPUTATIONS

VOLUME: 03 ISSUE: 03

RECEIVED: JULY 08 REVISED: AUGUST 04 ACCEPTED: AUGUST 24 PUBLISHED: SEPTEMBER 12

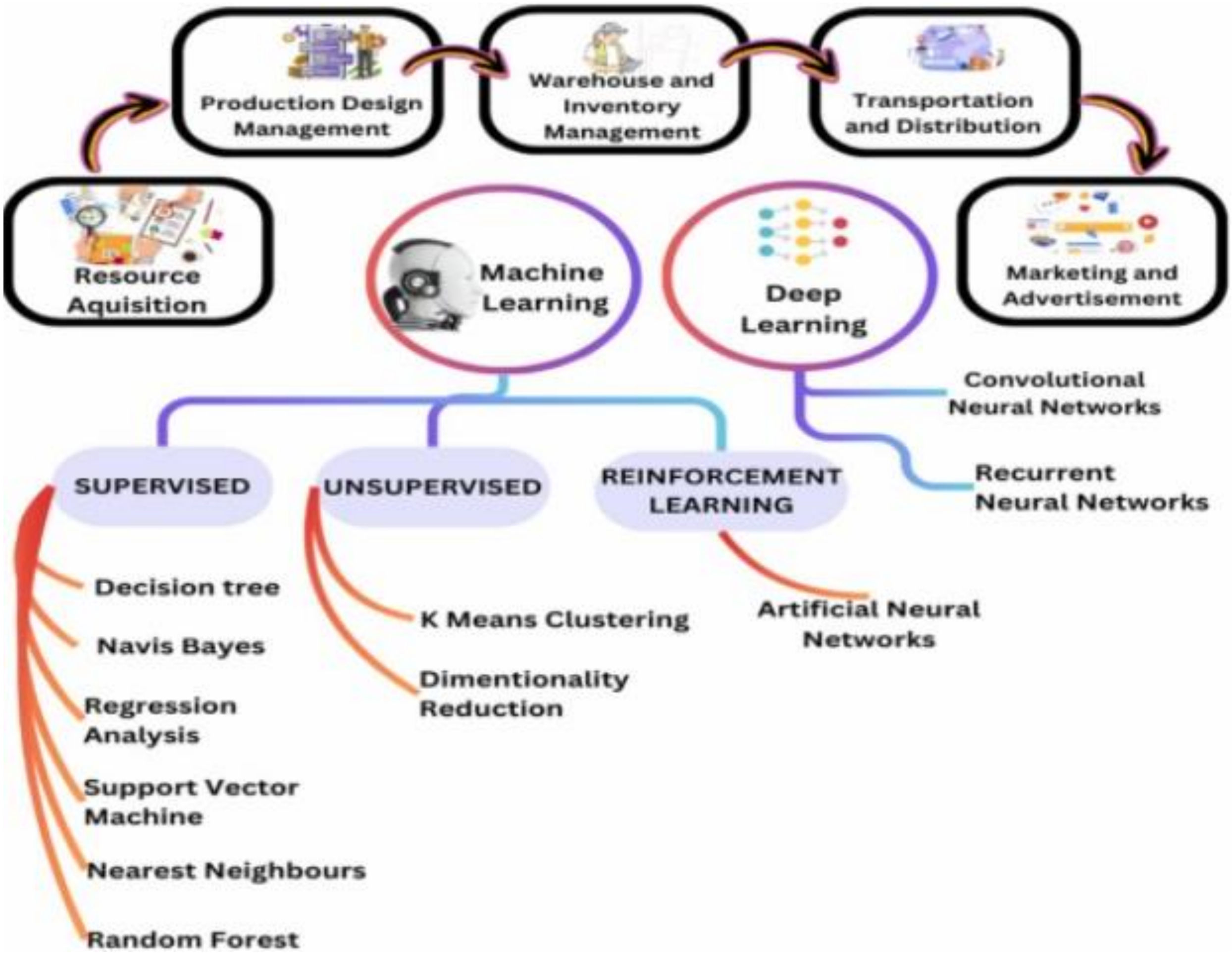


Fig 1: Enhancing supply chain management

1.1. Problem Statement and Research Questions

The complexity of global supply chains exposes them to external risks that challenge the accuracy of prediction. A literature survey in risk prediction reveals insufficient focus on technological risk from a supply chain management perspective. Numerous global supply chain risk indicators are available, but their accuracy has yet to be evaluated, even for demand risk. Machine learning, widely adopted in other sectors, requires Artificial Intelligence-enhanced Framework components and evaluation criteria reflecting the ultimate decision context to be directive for risk authorities. The underlying literature suggests three hypotheses targeting predictive performance from different angles: State-of-the-Art indicators focusing on Domain-Relevant signals are more Accurate than a Naïve-Estimate; Event-Owned signals possess predictive power; and Geopolitical, Meteorological, and Economic-Foundation factors improve prediction.

As the world heads deeper into the fourth industrial revolution, AI continues to expand its influence on supply chains. Worldwide supply chain difficulties, exacerbated by the COVID-19 pandemic, have introduced intense instability and uncertainty, underscoring the importance of supply chain resilience. Authorities are continuously exploring better means to predict supply

ISSN: 3067-4166

AMERICAN DATA SCIENCE JOURNAL FOR ADVANCED COMPUTATIONS

VOLUME: 03 ISSUE: 03

RECEIVED: JULY 08 REVISED: AUGUST 04 ACCEPTED: AUGUST 24 PUBLISHED: SEPTEMBER 12

chain risks in order to formulate corrective actions. These efforts are instrumental in fostering resilient supply chains and preventing future disasters.

Equation 1: Linear score → probability (sigmoid)

1. Linear score:

$$z_i = \beta_0 + \sum_{j=1}^{d} \beta_j x_{ij} = \beta_0 + \boldsymbol{\beta}^\top \mathbf{x}_i$$

2. Convert score to probability using the sigmoid:

$$\sigma(z) = \frac{1}{1 + e^{-z}}$$

So predicted probability:

$$\hat{p}_i = \sigma(z_i) = \frac{1}{1 + e^{-(\beta_0 + \boldsymbol{\beta}^\top \mathbf{x}_i)}}$$

1.2. Relevance to Global Supply Chains

Risk prediction in data engineering for global supply chains concerns the assessment of the external environment and its disruption potential—factors in demand and supply chains that lie outside the company's borders. Relevant operational corridors are determined by the point of risk measurement and the prediction horizon. In addition to logistics, demand, and supply-related risk signals, others emerging in the political, climate, or economic space require focused, continuous monitoring to help business leaders anticipate global developments that might jeopardize their ability to deliver products or services on time and at the right quality.

To be effective, risk signals must be cognizant of the decision-making context and must use a range of data sources, both structured and unstructured, covering the territories from which goods and services are sourced or supplied. Geopolitical and economic factors related to the territory must, therefore, be continuously probed (rich data sources on nations are provided for instance by the World Bank, OECD, and UN), and weather-related events must be tracked for the entire product lifecycle, with a minimum focus on loading, shipping, receiving, and delivery days. Health-related risks are also increasingly relevant during pandemic or epidemic periods—and great pandemic risk datasets are available. Key risk indicators, their causal rationale, and suggested measurement approaches are listed.

2. Theoretical Foundations

Effective and trusted prediction of risks in global supply chains stems from suitable data engineering and AI-enhanced, machine-learning-based models. For risk-aware analytics, the creation and operation of robust data pipelines must-as a foundation-be established by designing the appropriate data-scientific-data-engineering modules. This includes the design of systems to initiate seamless streams of real-time data and event-driven processing, source risk signals from data of heterogeneous nature as well as identify and integrate latent risk signal components delivered by neural networks.

ISSN: 3067-4166
AMERICAN DATA SCIENCE JOURNAL FOR ADVANCED COMPUTATIONS
VOLUME: 03 ISSUE: 03
RECEIVED: JULY 08 REVISED: AUGUST 04 ACCEPTED: AUGUST 24 PUBLISHED: SEPTEMBER 12

The inference of core demand, supply and logistics risk indicators must be pursued alongside those of external risk-related factors, including geopolitical-shock intensity estimates, the probability of extraordinary weather events, the determination of energy price shocks, a measurement of future economic contraction, etc. Predictive performance metrics suited to measuring the impact of the prediction on risk-aware decisions must be specified. Such a modelling framework would be of immediate and future value to risk-conscious decision-makers.

2.1. Data Engineering for Risk Analytics

Well-defined data pipelines and data schemas, a robust feature store, and comprehensive lineage details are essential for analytic applications in demand, supply, and logistic risk prediction. Data scientists can focus on model development and evaluation only if clean training data is available in the correct format.

Data ingestion is the first step in the data preparation process. The aim is to collect the data from various sources and provision the data for analytical processing. Data is typically ingested from external, often publicly available sources, as well as internal data repositories. The motive can vary depending on the data source in question. Data produced inside an organization is ingested for reactive, historical analytics, while data stream-ing from social media and other real-time feeds is ingested for proactive, forward-looking analytics. Data is often generated in a continuous fashion and requires a streaming ingestion backbone as well as accompanying infrastructure for event-driven processing. Pointing to past data to identify possible risk events or conducting a periodic analysis reruns the complete data ingestion architecture using a batch-oriented back-end.

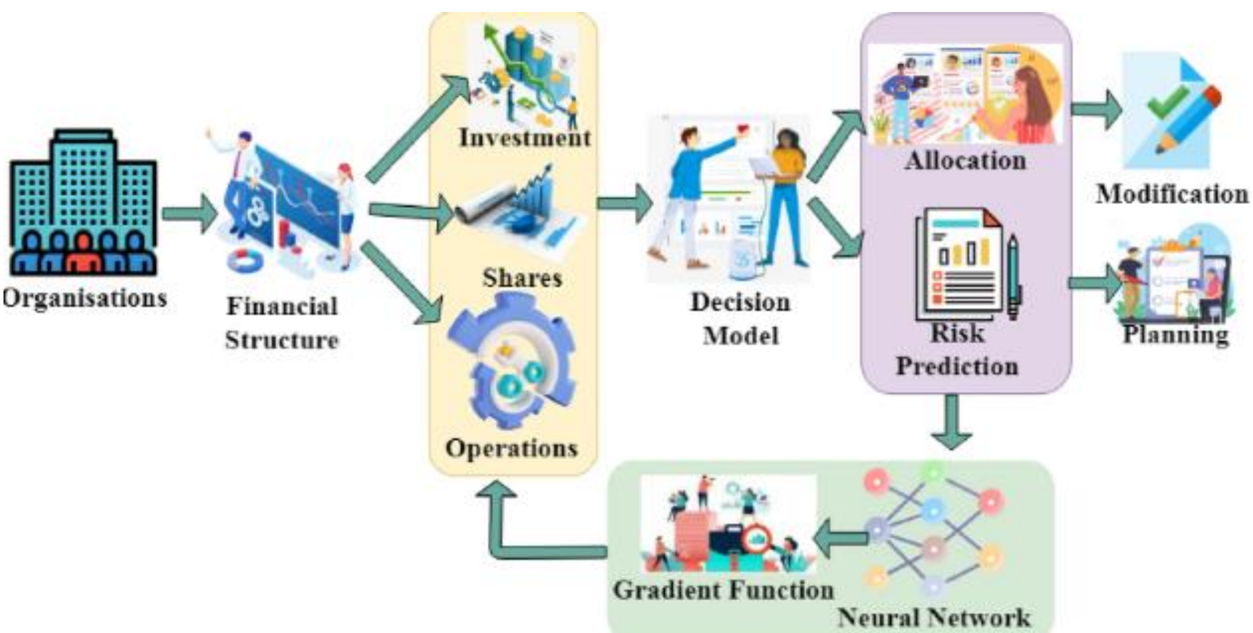


Fig 2: Data Engineering for Risk Analytics

2.2. Artificial Intelligence and Machine Learning in Risk Prediction

Well-suited models for risk prediction are interpretable, either using a traditional classification methodology with logistic regression and relevant K-level data or PH, or probabilistic models capable of producing reasonably accurate risk signals without a large number of positive cases for the target events. Machine learning methods are also appropriate within an AI-enhanced Data

ISSN: 3067-4166

AMERICAN DATA SCIENCE JOURNAL FOR ADVANCED COMPUTATIONS

VOLUME: 03 ISSUE: 03

RECEIVED: JULY 08 REVISED: AUGUST 04 ACCEPTED: AUGUST 24 PUBLISHED: SEPTEMBER 12

Engineering framework when more complex relationships are expected or when sufficient positive samples exist, utilizing either supervised or unsupervised learning. Classification problems can be approached expertially or purely data-driven, with classical ML methods remaining suitable within a small domain. Within a wider context, however, including varying types of risks, ensemble learning—such as stacking—offers improved predictive accuracy alongside a lower risk of overfitting, easily allowing gradient boosting. Possible unsupervised learning tasks also include assessment of multimodal signals to identify domain-independent precursors of forthcoming risks within the other domains.

Depending on the type of risk predicted and underlying domain knowledge, large language models (LLMs) may also play a useful role. For example, in predicting the next month's demand, the historical patterns obvious from time series can be enough, even without additional real-time prompts. When predicting supply or logistic disruptions, however, decision-making behaviour is increasingly important, and LLMs could help construct additional variables from recent news articles, statements, and similar data to complement the pure time series information. When turning towards the next several months, all three domains offer a less-reliable but deep contextualisation. Factors that may not be quantifiable can still have great importance, and real-world human intuition assesses such nuances polydimensionally much better than unidirectional time series extrapolation permits. For such requirements, recent huge models like ChatGPT represent a welcome advance. Though still in experimental stages and even requiring true prompts within the built platforms, they indeed offer significant potential for carefully integrated natural-language augmentation of more traditional prediction frameworks, especially in generating input variables from rich but less-structured information sources.

Equation 2: Likelihood and log-likelihood (step-by-step)

Given $\hat{p}_i$, the Bernoulli probability of observing $y_i$is:

$$P(y_i \mid \mathbf{x}_i) = \hat{p}_i^{y_i}(1-\hat{p}_i)^{1-y_i}$$

Assuming i.i.d. across $i = 1, \dots, n$, likelihood:

$$L(\boldsymbol{\beta}) = \prod_{i=1}^{n} \hat{p}_i^{y_i}(1-\hat{p}_i)^{1-y_i}$$

Log-likelihood:

$$\ell(\boldsymbol{\beta}) = \log L = \sum_{i=1}^{n} (y_i \log \hat{p}_i + (1-y_i)\log(1-\hat{p}_i))$$

Negative log-likelihood (a common loss a.k.a. log loss / cross-entropy):

$$\mathcal{L}(\boldsymbol{\beta}) = -\ell(\boldsymbol{\beta})$$

3. Architectural Paradigms

Architects can turn to data engineering frameworks that enable the ingestion and integration of risk data from public and private systems across different jurisdictions and countries, that establish quality and governance standards for those data, and that permit them to be formally associated with the regions and actors that they affect. The definition of data flows and events can

ISSN: 3067-4166

AMERICAN DATA SCIENCE JOURNAL FOR ADVANCED COMPUTATIONS

VOLUME: 03 ISSUE: 03

RECEIVED: JULY 08 REVISED: AUGUST 04 ACCEPTED: AUGUST 24 PUBLISHED: SEPTEMBER 12

then be complemented with attention to real-time streams, AI-enhanced multimodal data fusion, and risk indicators built from supply chain demand, supply, logistics, geopolitical, weather, and economic risks. Supply chain professionals charged with risk assessment can besides seek risk predictors drawn from these frameworks, with the metric system evaluating the contribution of the predictors to subsequent business decisions and the build quality of the underlying models. Road testing and validation of the deployed predictive models can assure the robustness of the ensuing risk framework.

All sources that produce data relevant for supply chain monitoring and risk assessment pose ingest and integration challenges. Data can flow from physical or cyber events, news wires, or social media streams and be consumed continuously and within very tight latency limits. Data for longer-term prediction horizons may be sourced from controlled datasets and historic archives, typically via batch ingestion. Regardless of production mode and timeline, partners from different organisational and governance domains will look for means of sharing data without opening sensitive private or strategic information. Reliable exchange relies on formally specified and machine-readable descriptions of the data, their contents, and their semantics, and on the availability of both country domain-specific schemas and numerous country-domain mapping definitions.

3.1. Data Ingestion and Integration in Global Networks

Data ingestion and integration cover the sources of data, de facto streams in supply chains, implications of batch versus real-time approaches, requirements for hybridity and interoperability, and challenges of cross-border interconnection. Supply chains process data from virtually all global information systems. Any significant event generates data that can be leveraged for risk assessment, so the central question is whether these activity streams can be fused in a coherent way. Data from supply chains, especially demand data, is often publicly unavailable; however, large corporations and dedicated start-ups offer indicators that reflect emerging shifts and risks at a global level.

Data streams in supply chain networks are largely unidirectional. During normal periods, individual demand, supply, manufacturing, transportation, and stock management processes generate large volumes of data. Although these processes are independent within a single company, they are not isolated in global networks. Disruption events introduce a second type of data structure: signals from external sources that impact supply chains. Different types of events alter these unidirectional streams, with demand usually leading supply, logistics, and manufacturing. In practice, these unidirectional data streams are complemented by bulk data transfer through the encrypted Internet. These Internet flows include the whole network of secure/unknown type flows plus traditional batch transfers, but they seldom contain classified data. Streaming data can be divided into two types: event-driven real-time streaming processes and repetitive batch-type processing. Event-driven streams are normally handled by specific engines, while repetitive processing can be built with lower-latency targets.

ISSN: 3067-4166

AMERICAN DATA SCIENCE JOURNAL FOR ADVANCED COMPUTATIONS

VOLUME: 03 ISSUE: 03

RECEIVED: JULY 08 REVISED: AUGUST 04 ACCEPTED: AUGUST 24 PUBLISHED: SEPTEMBER 12

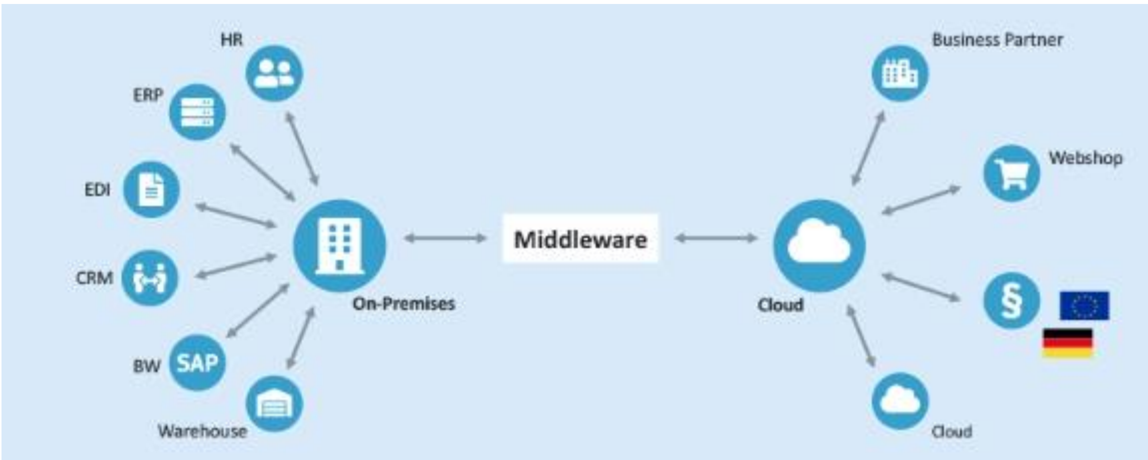


Fig 3: Data Ingestion vs Data Integration

3.2. Data Quality, Governance, and Stewardship

Data integration across countries, industries, organizations, and disparate partners demands stringent data quality standards. Such standards are often spelled out in terms of the data quality dimensions proposed by Wang and Strong, which include accuracy, completeness, consistency, timeliness, credibility, and ease of use. Quality metrics require the specification of thresholds, rules or heuristics and, if applicable, source-provided ground truth. Metadata should specify the actors responsible for assessing and certifying data quality. Governance processes, including vocabularies and ontologies, must address and manage the inevitable imperfections in the data in terms of completeness, correctness, and bias introduced by missing, conflicting, inaccurate or untrustworthy data.

Many of these features fall under the domain of data management and stewardship. Obviously, owners of private data have the prerogative of deciding how and when these data can be used by others. Such stewardship controls may include procedures to ensure that only authorized actors are applying the data in appropriate contexts, and auditing capabilities to check compliance. Economic issues associated with the sharing of private data also fall into the stewardship area, and agreements to compensate data owners or to share cloud costs can be implemented. These policies may even extend to providing actual compensation in the event of a data breach or infrastructure activity that causes severe damage in the target localities.

Equation 3: Confusion matrix and "accuracy" metrics (derived)

Define counts:

- TP = true positives = $\sum \mathbf{1}(\hat{y}_i = 1, y_i = 1)$
- FP = false positives = $\sum \mathbf{1}(\hat{y}_i = 1, y_i = 0)$
- TN = true negatives = $\sum \mathbf{1}(\hat{y}_i = 0, y_i = 0)$
- FN = false negatives = $\sum \mathbf{1}(\hat{y}_i = 0, y_i = 1)$

4.1 Accuracy

Step-by-step:

ISSN: 3067-4166
AMERICAN DATA SCIENCE JOURNAL FOR ADVANCED COMPUTATIONS
VOLUME: 03 ISSUE: 03
RECEIVED: JULY 08 REVISED: AUGUST 04 ACCEPTED: AUGUST 24 PUBLISHED: SEPTEMBER 12

- Correct predictions = TP + TN
- Total = TP + FP + TN + FN = n

So:

$$\text{Accuracy} = \frac{TP + TN}{n}$$

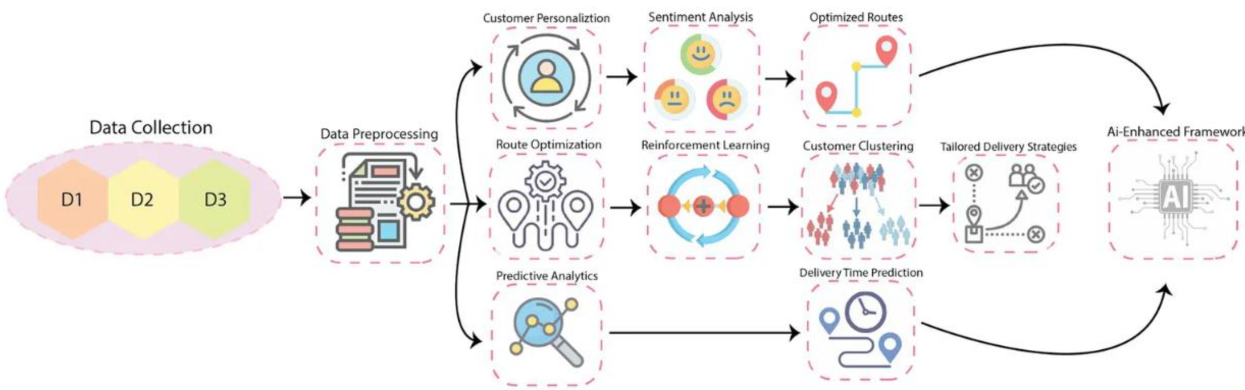
4. AI-Enhanced Framework Components

Supported by the data architecture described in the previous section, AI-enhanced frameworks for predicting supply chain risk include two essential components. The first is the ability to reliably ingest and rapidly process real-time data streams. The second is the capability to fuse information from dissimilar sources and modalities—spatial and temporal, carried by text, voice, image, or directly connected sensors—into coherent signals that disclose key changes in global risk: aggravated demand imbalance, supply shortage or disruption, and transport bottleneck.

4.1. Real-Time Data Streams and Event-Driven Processing

Responsiveness depends on the ability to react to events as they occur, often long before they materialize as changes in quantitative measures of risk. Many such events play out in the geopolitical, military, and economic domains, and they are traditionally monitored by analysts in think tanks, government agencies, and corporations. As their impact on supply risk becomes apparent, more generic quantifiable indicators are moved into the foreground. The de facto latencies associated with event monitoring, detection, and warning are already defined, as are the signals, alerts, and synopses needed by practitioners to facilitate response.

Incorporating these requirements into a data system architecture enables the development of event schemas whose completion is supported by pipelines for event detection, monitoring, and alerting. These schemas are then naturally filled at run time by processing frameworks offering the necessary speed, precision, and volume. Latency targets imposed by industry stakeholders, particularly manufacturers of complex products with sensitive supply chains, determine the actual measured latencies but also influence design decisions. Stream processing frameworks such as Apache Flink are naturally suited to the low-latency pipe and filter structures required. Low-latency alerting services fed by many of these streams, including signals flagged by the structures themselves, are therefore well supported.



ISSN: 3067-4166
AMERICAN DATA SCIENCE JOURNAL FOR ADVANCED COMPUTATIONS
VOLUME: 03 ISSUE: 03
RECEIVED: JULY 08 REVISED: AUGUST 04 ACCEPTED: AUGUST 24 PUBLISHED: SEPTEMBER 12

Fig 4: AI-Enhanced Framework Components

4.1. Real-Time Data Streams and Event-Driven Processing

Event-driven architectures couple event schemas with low-latency stream-processing engines that emit alerts on risk signals. Speed throughput meets the requirements for pipeline latency, which varies from seconds for social media data to hours or weeks for economic reports. Advanced message queue services publish and process events from global supply chains. Leading cloud providers embed specialized data-processing engines that support diverse stream-processing recipes, from simple alerting to data enrichment. Complex events combine multiple sources, clustering, and multimodal data fusion.

Events appear to service the downstream data pipeline whenever processing latency allows. Temporal and probabilistic data fusion at the pipe output reconciles conflicting or incomplete signals: when estimating risk exposure, a lack of or minor uncertainty in one source acquires lower weight than equivocal or contradictory inputs. We should avoid exposing technical overhead to the final application. Alerts require warning in advance of possible action, such as reallocation or material substitution, decision-making input, or mitigation expectations attachment.

Equation 4: Definition (binning-based reliability)

1. Partition predicted probabilities into bins (e.g., 0–0.1, 0.1–0.2, …).
2. For a bin B_k, compute:

Average predicted risk:

$$\bar{p}_k = \frac{1}{|B_k|} \sum_{i \in B_k} \hat{p}_i$$

Empirical event rate:

$$\bar{y}_k = \frac{1}{|B_k|} \sum_{i \in B_k} y_i$$

A perfectly calibrated model has $\bar{y}_k \approx \bar{p}_k$ for all k.

4.2. Multimodal Data Fusion for Risk Signals

AI-enhanced data engineering frameworks for predicting global supply chain risks provide essential ingredients for uncovering demand, supply, and logistics risk signals. A number of additional components, however, help complete the picture. Real-time data streams minimize latencies from data generation to delivery. Event-driven processing combines low-latency updates across sources and streaming pipelines with infrequent, more involved batch jobs, ensuring timely action without compromising detection quality. Multimodal data fusion techniques align different types of distinct but related risk signals and synthesize models that receive inputs from mutually supportive streams, catering to those cases where some signals are unavailable or conflicting.

The global economy is walking on broken glass and trying to pick up the pieces without losing balance. Worldwide demand and supply risks have trickled in, but logistics have emerged as the biggest Achilles heel, with serious ramifications for the overall economy. New sources of risk are bubbling to the surface, adding new challenges to an already fragile supply-demand balance.

ISSN: 3067-4166
AMERICAN DATA SCIENCE JOURNAL FOR ADVANCED COMPUTATIONS
VOLUME: 03 ISSUE: 03
RECEIVED: JULY 08 REVISED: AUGUST 04 ACCEPTED: AUGUST 24 PUBLISHED: SEPTEMBER 12

Geopolitical uncertainty, severe weather events, and a potentially long recession are just three of the ever-growing set of threats that affect the economy in real time. Signalling from such factors must therefore be considered as well.

5. Risk Prediction in Practice

Supply chain risk prediction remains a challenging task despite significant advances in demand forecasting and machine learning. Risk indicators can predict the safety of the supply chain network and serve as early warning signals for supply disruptions. Predictive values do not need to be high; they should, however, be sufficiently informative to determine supply chain resilience and enable preparedness to withstand high-risk conditions. The main risk factors associated with demand, supply, and logistics risk are identified, along with causal explanations and practical approaches for risk signal development.

Risk prediction is further complicated by the external factors shaping the global supply chain landscape. Real-time geopolitical developments, severe weather events, transportation and logistic disruptions, economic distress, the COVID-19 pandemic, and the protracted war in Ukraine all pose significant supply chain risks. While these forces are external to the supply chain as a network of companies, they cross borders and can trigger a sequence of events affecting the risk profile of the supply chain network. Hence, predictive models of supply chain risk should draw on these external risk factors. Major risk factors are identified alongside relevant data sources, optimal data change frequencies, and the implications for integration into predictive models.

5.1. Demand, Supply, and Logistics Risk Indicators

Global supply chains continuously adapt to external pressures, and their flexibility can determine success. Demand-side shocks are often pivotal but can remain hidden for long periods. By contrast, supply- and logistics-side shocks tend to become evident earlier but are also more common. Still, predictive models for demand, supply, and logistics disruptions remain scarce.

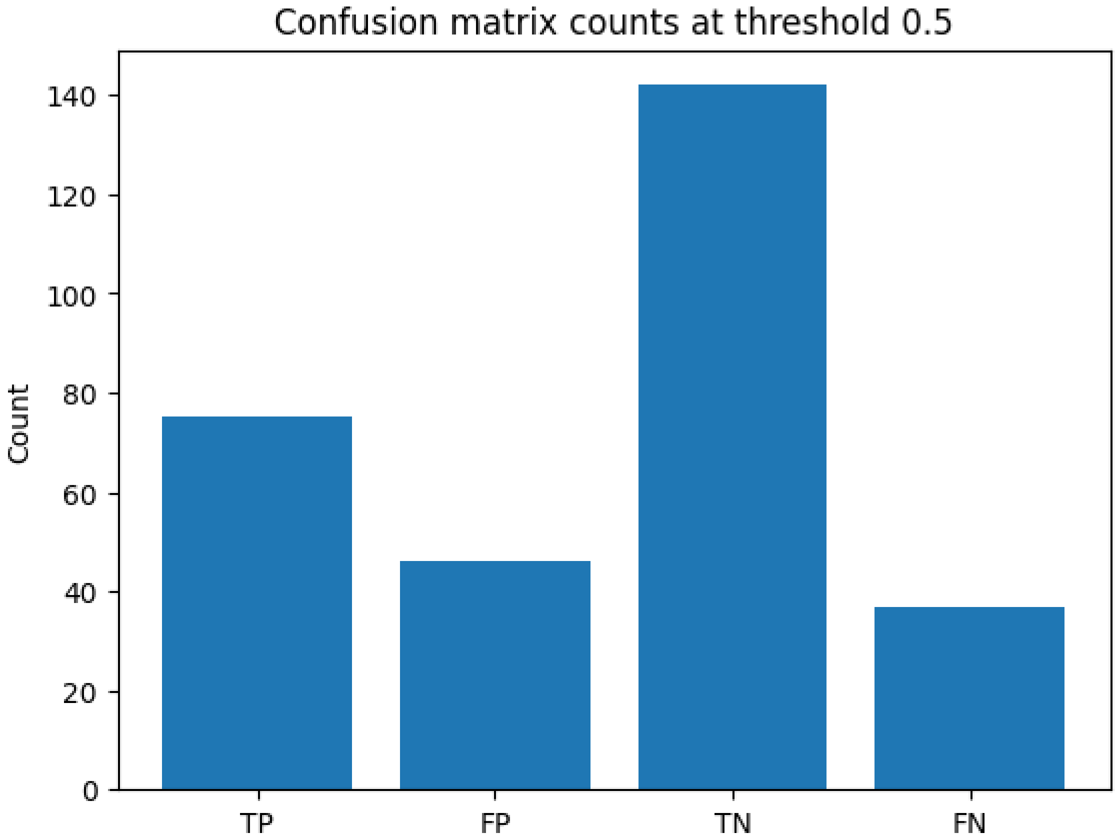
Core indicators, causal rationales, and feasible measurement approaches are summarized. Typically, demand disruptions are related to fast-growing subclasses (e.g. fashion-clothing) and persistent-worldwide shocks (e.g. COVID). Supply risks include capacity constraints at suppliers, component shortages, and bottlenecks on adjacent links. Measures for logistics risks account for congestion levels at major ports and airports, lifting costs, and transit times.

ISSN: 3067-4166

AMERICAN DATA SCIENCE JOURNAL FOR ADVANCED COMPUTATIONS

VOLUME: 03 ISSUE: 03

RECEIVED: JULY 08 REVISED: AUGUST 04 ACCEPTED: AUGUST 24 PUBLISHED: SEPTEMBER 12



5.2. Geopolitical, Weather, and Economic Risk Factors

Global supply chains operate in an environment of continuous change, including social, political, and weather events that can affect demand, production, logistics, and supply continuity. Within this context, risk indicators must integrate data representing holiday celebrations in various countries, geopolitical tensions, adverse weather conditions, economic developments, and macroeconomic variables such as GDP, inflation rates, and interest rates.

The following list summarizes such risk factors with data sources and measurement:

1. **Geopolitical risks.** Indicators of geopolitical risk can be derived from an innovative database that compiles information on various conflicts worldwide, including those involving major powers. The signal is constructed using a binary indicator that takes value 1 when a conflict of significant intensity occurs within the region of the focal node. The dataset covers several decades and is frequently updated to capture changes in the geopolitical landscape. Additional signals can also be fed into a conflict index, which assigns weights to different countries or regions according to their economic importance.

2. **Weather.** Daily weather information may be retrieved from sources such as the National Oceanic and Atmospheric Administration's Global Historical Climatology Network, which collects ground station data from around the world. The major weather conditions affecting risk levels are adverse temperature, precipitation, and wind events. The signals for wind storms,

ISSN: 3067-4166
AMERICAN DATA SCIENCE JOURNAL FOR ADVANCED COMPUTATIONS
VOLUME: 03 ISSUE: 03
RECEIVED: JULY 08 REVISED: AUGUST 04 ACCEPTED: AUGUST 24 PUBLISHED: SEPTEMBER 12

floods, snow, and freezing can be obtained using a binary variable that takes value 1 when exceeding a certain threshold. The temperature and precipitation signals may be built using a simple aggregation between all countries/regions directly linked in the network and considering one single country/region.

3. **Economic development.** Macroeconomic news, including GDP growth forecasts, inflation levels, and interest rates, can be collected from central banks and other related institutions. Such information can be consumed by the risk prediction system either as a streaming feed when released or through an aggregated database containing all relevant information over the past months. In addition, qualitative assessments of a wide set of countries are published regularly by the Economist Intelligence Unit and similar databases.

Equation 5: ROC curve and AUC (explicitly mentioned)

For a threshold t:

- True Positive Rate (TPR) = Recall:

$$TPR(t) = \frac{TP(t)}{TP(t) + FN(t)}$$

- False Positive Rate (FPR):

$$FPR(t) = \frac{FP(t)}{FP(t) + TN(t)}$$

The ROC curve plots $TPR(t)$vs $FPR(t)$as $t$sweeps from 1 down to 0.

Mathematically:

$$AUC = \int_0^1 TPR(FPR)\, d(FPR)$$

Numerically (trapezoidal rule). If points are (FPR_k, TPR_k)sorted by FPR:

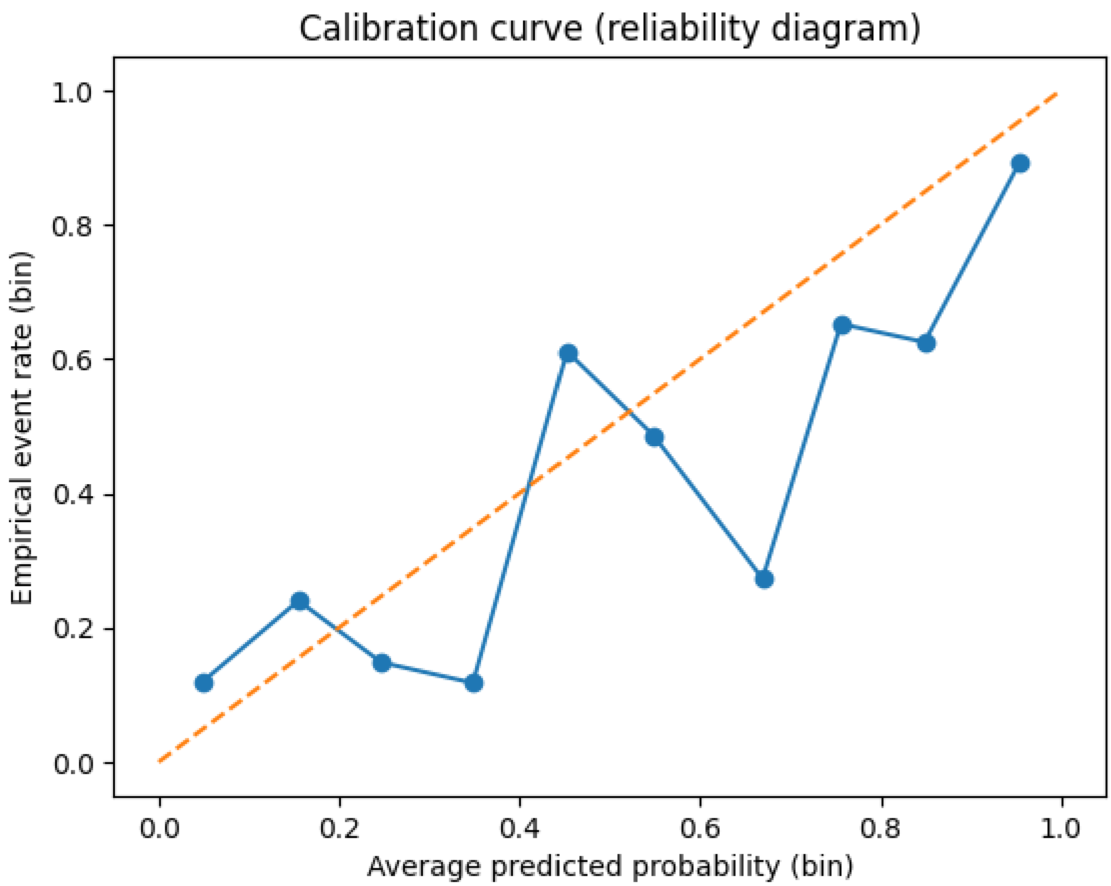
$$AUC \approx \sum_{k=1}^{m-1} (FPR_{k+1} - FPR_k) \cdot \frac{TPR_{k+1} + TPR_k}{2}$$

6. Evaluation and Validation

Every prediction task requires metrics to assess predictive performance. In risk prediction, the precise nature of these metrics depends on who uses the indicators and actions taken in response to certain events. For instance, risk thresholds informed by demand forecasts or supply chain performance measures could differ according to the decision-making context. However, accuracy and calibration are always relevant. Several decision-analytic measures applied combine prediction accuracy and uncertainty quantification.

ISSN: 3067-4166
AMERICAN DATA SCIENCE JOURNAL FOR ADVANCED COMPUTATIONS
VOLUME: 03 ISSUE: 03
RECEIVED: JULY 08 REVISED: AUGUST 04 ACCEPTED: AUGUST 24 PUBLISHED: SEPTEMBER 12

Robustness, generalization, and calibration remain among the most critical factors when applying a model to a new domain, such as weather shocks, to related but different data sources, like economic indicators, and across time. Such tests are crucial to establishing the credibility of a complex supply chain model that ultimately depends on event occurrence to validate future predictions.



6.1. Metrics for Predictive Performance

Evaluation of prediction performance is often context-dependent. When firms maintain specific risk thresholds and use risk predictions primarily to guide contingency responses, standard accuracy measures assume less importance. In such cases, failures to predict a significant risk may be more costly than false alarms. Calibration, which quantifies the relation between realized and predicted risk levels, thus becomes more salient. Decision-analytic performance measures such as the area under the receiver operating characteristic curve (AUC) or the area under the precision-recall curve (AUC/PR) weigh costs and benefits of predictions according to the potential impact of risk prediction decisions. AUCs provide an overall quality measure that considers how predictions influence firm decision-making.

Nevertheless, firms often seek predictive features to drive adjustments in strategy, capacity planning, and operational execution in both the near and longer terms. For such applications, metrics that directly evaluate predictive accuracy and robustness remain

ISSN: 3067-4166

AMERICAN DATA SCIENCE JOURNAL FOR ADVANCED COMPUTATIONS

VOLUME: 03 ISSUE: 03

RECEIVED: JULY 08 REVISED: AUGUST 04 ACCEPTED: AUGUST 24 PUBLISHED: SEPTEMBER 12

essential during model design. Moreover, the calibration need not be dominated by relatively simple yet valid predictive features; improving quantification of less frequent but more consequential failures is also of interest.

6.2. Robustness, Generalization, and Calibration

Prediction-driven decisions often involve high stakes, and validation must therefore ensure that predictive models can accurately identify risk events outside the training window. This requires some degree of cross-domain testing, as supply chain disruptions present their own set of 'known knowns and unknowns' [44]. Events of rare and extreme severity—such as major earthquakes, terrorist attacks, and pandemic outbreaks—deserve special attention, particularly under regression tasks, for which detection failure can signal catastrophic miscalibration [45]. Stress tests that adjust risk-feature mappings beyond their training ranges can further help identify model limitations and develop an understanding of domain dependencies.

Incorporation of external risk factors from a different domain creates generalization challenges, but the importance of these signals—such as the Russia-Ukraine conflict or the COVID pandemic—for risk prediction remains undeniable. Domain adaptation strategies can enable the reuse of successively trained models in different supply chain contexts. Approaches exploiting information in model ensembles for known-unknown assessments can also systematically quantify the uncertainty associated with predictions currently beyond the system's experience. Finally, the assessment and communication of uncertainty through such uncertainty quantification techniques can significantly enhance the practical value of predictive models.

7. Conclusion

Contributions, limitations, and future research directions are summarized in this conclusion. The proposed AI-enhanced data engineering frameworks address critical yet generally underserved aspects of supply chain risk prediction. Data pipelines tailored to risk analytics support the ingestion, transformation, and storage of assorted data sources in anticipation of data-hungry AI/ML models. Indeed, the construction and management of feature stores are essential to supply global-risk predictions with requisite real-world signals. Evolving demand, supply, and logistics patterns are captured through well-timed, event-driven data streams; multimodal data-fusion approaches identify impending threats not evident in any individual stream. Finally, predictive performance of the risk models—or lack thereof—is comprehensively evaluated with respect to accuracy, generalization, calibration, and robustness.

Despite these contributions, important risk-prediction elements remain neglected. Internal supply chain disruptions have yet to be embedded in the framework, with changes in need priority order still requiring attention. External risk factors beyond geopolitical threats—such as weather, economic, and environmental conditions—warrant further exploration. Nevertheless, the integration of geopolitical, weather, and economic aspects should ultimately enhance risk predictions in terms of accuracy, robustness, and calibration. Future research might also consider issues of governance and acceptance, charting the territory between widespread embarrassment and credible prediction.

bin	count	avg_pred	emp_rate
1	67	0.04867194753294893	0.11940298507462686
2	50	0.15507903005871648	0.24
3	27	0.24695906134954068	0.14814814814814814

ISSN: 3067-4166
AMERICAN DATA SCIENCE JOURNAL FOR ADVANCED COMPUTATIONS
VOLUME: 03 ISSUE: 03
RECEIVED: JULY 08 REVISED: AUGUST 04 ACCEPTED: AUGUST 24 PUBLISHED: SEPTEMBER 12

bin	count	avg_pred	emp_rate
4	17	0.348410253463033	0.11764705882352941

Table: Calibration table (10 bins)

7.1. Future Directions

AI-enhanced supply chain risk prediction is in its infancy, yet many avenues are ripe for exploration. The focus can now shift toward building, testing, and deploying frameworks based on the theory and considerations presented here. Empirical studies should fill critical gaps in the understanding of risk signals, shape predictive models for sentinel indicators, and develop frameworks that integrate and validate risk-prediction models. In parallel, the convergence of two more general trends—growing amounts of real-time data in supply chains and the need for risk-aware operations—offers fertile ground for engendering real-time, risk-stage-data-centric microservices for supply chain operations. Here, immature developments in the combination of risk signals should gradually culminate into a common risk status for any given shipment.

Another dimension worth considering relates to economic inequalities, economic hierarchies, environmental hostilities, and the power dynamics amidst social groups. Once nodes and flows are mapped based on voice-vote, responsibility, and economic position, risk signals can be robustly developed and deployed, since behaving as a bigger-than-self local government is crucial for natural disaster. Group of companies forming local partnerships can be treated as a unit to monitor risks since these partnerships should be formed to survive in normal situations with a bigger-than-self motivation.

8. References

1. Inala, R. (2025). A Unified Framework for Agentic AI and Data Products: Enhancing Cloud, Big Data, and Machine Learning in Supply Chain, Insurance, Retail, and Manufacturing. EKSPLORIUM-BULETIN PUSAT TEKNOLOGI BAHAN GALIAN NUKLIR, 46(1), 1614-1628.
2. Huang, S. (2025). AI-driven early warning systems for supply chain risk detection: A machine learning approach. Academic Journal of Computing & Information Science, 8(9), 92-107. https://doi.org/10.25236/AJCIS.2025.080913.
3. Garapati, R. S., & Daram, D. S. B. (2025). AI-Enabled Predictive Maintenance Framework For Connected Vehicles Using Cloud-Based Web Interfaces. Available at SSRN 5524261.
4. Kephart, J. O., & Chess, D. M. (2003). The vision of autonomic computing. Computer, 36(1), 41–50.
5. Aitha, A. R., & Jyothi Babu, D. A. (2025). Agentic AI-Powered Claims Intelligence: A Deep Learning Framework for Automating Workers Compensation Claim Processing Using Generative AI. Available at SSRN 5505223.

ISSN: 3067-4166

AMERICAN DATA SCIENCE JOURNAL FOR ADVANCED COMPUTATIONS

VOLUME: 03 ISSUE: 03

RECEIVED: JULY 08 REVISED: AUGUST 04 ACCEPTED: AUGUST 24 PUBLISHED: SEPTEMBER 12

6. Jahin, S. (2024). Intelligent supply chain management: A systematic literature review on artificial intelligence contributions. Information, 16(5), 399-420. https://doi.org/10.3390/info16050399.
7. Nagabhyru, K. C., & Babu, A. J. Human In The Loop Generative AI: Redefining Collaborative Data Engineering For High Stakes Industries.
8. Erl, T. (2016). Microservices design patterns. Prentice Hall.
9. Gottimukkala, V. R. R. (2025). Generative AI for Exceptions and Investigations: Streamlining Resolution Across Global Payment Systems. Journal of International Commercial Law and Technology, 6(1), 969-972.
10. Fowler, M. (2018). Refactoring (2nd ed.). Addison-Wesley.
11. Segireddy, A. R. (2025). GENERATIVE AI FOR SECURE RELEASE ENGINEERING IN GLOBAL PAYMENT NETWORK. Lex Localis: Journal of Local Self-Government, 23.
12. Gamma, E., Helm, R., Johnson, R., & Vlissides, J. (1994). Design patterns. Addison-Wesley.
13. Amistapuram, K. (2025). GENERATIVE AI FOR CLAIMS EXCEPTIONS AND INVESTIGATIONS: ENHANCING RESOLUTION EFFICIENCY IN COMPLEX INSURANCE PROCESSES. Available at SSRN 5785482.
14. Zaharia, M., et al. (2010). Spark: Cluster computing with working sets. HotCloud.
15. Sheelam, G. K. (2025). Agentic AI in 6G: Revolutionizing Intelligent Wireless Systems through Advanced Semiconductor Technologies. Advances in Consumer Research.
16. Lakshman, A., & Malik, P. (2010). Cassandra. ACM SIGOPS Operating Systems Review, 44(2), 35–40.
17. Nagubandi, A. R. (2025). PIONEERING SELF-ADAPTIVE AI ORCHESTRATION ENGINES FOR REAL-TIME END-TO-END MULTI-COUNTERPARTY DERIVATIVES, COLLATERAL, AND ACCOUNTING AUTOMATION: INTELLIGENCE-DRIVEN WORKFLOW COORDINATION AT ENTERPRISE SCALE. Lex Localis, 23(S6), 8598-8610.
18. Stonebraker, M., & Çetintemel, U. (2005). One size fits all? ICDE Proceedings, 2–11.
19. Yandamuri, U. S. (2022). Big Data Pipelines for Cross-Domain Decision Support: A Cloud-Centric Approach. International Journal of Scientific Research and Modern Technology, 227.
20. Moreira, M. W. L., et al. (2018). IoT-based smart healthcare systems. Sensors, 18(4), 1155.
21. Sasi Kumar Kolla. (2023). Big Data–Driven Machine Learning Frameworks for Clinical Risk Prediction. International Journal of Medical Toxicology and Legal Medicine, 26(3 and 4), 44–59. Retrieved from https://ijmtlm.org/index.php/journal/article/view/1456
22. Sriram, H. K., Gadi, A. L., & Challa, K. (2025). Leveraging AI, ML, and Gen AI in Automotive and Financial Services: Data-Driven Approaches to Insurance, Payments, Identity Protection, and Sustainable Innovation. Anil Lokesh and Challa, Kishore and

ISSN: 3067-4166

AMERICAN DATA SCIENCE JOURNAL FOR ADVANCED COMPUTATIONS

VOLUME: 03 ISSUE: 03

RECEIVED: JULY 08 REVISED: AUGUST 04 ACCEPTED: AUGUST 24 PUBLISHED: SEPTEMBER 12

singreddy, Sneha, Leveraging AI, ML, and Gen AI in Automotive and Financial Services: Data-Driven Approaches to Insurance, Payments, Identity Protection, and Sustainable Innovation (March 25, 2025).

23. Kumar, K. M., Parasar, A., Walia, A., Inala, R., & Thulasimani, T. (2025, August). Enhancing Risk Management Strategies in Financial Institutions Using CNN and Support Vector Regression. In 2025 5th Asian Conference on Innovation in Technology (ASIANCON) (pp. 1-6). IEEE.
24. Koller, D., & Friedman, N. (2009). Probabilistic graphical models. MIT Press.
25. Rao, A. N., Garapati, R. S., Suganya, R. T., Kaliappan, A., & Kamaleshwar, T. (2025, August). Smart Solar Harvesting and Power Management in IoT Nodes Through Deep Learning Models. In 2025 2nd International Conference on Intelligent Algorithms for Computational Intelligence Systems (IACIS) (pp. 1-6). IEEE.
26. Liu, F., et al. (2025). Foundational architecture for AI agents in healthcare. Cell Reports Medicine, 6(10), 102374.
27. Paleti, S., Baliyan, M., Aitha, A. R., Reddy, B. A., Bhadauria, G. S., & Sing, S. A. (2025, August). Graph—LSTM Hybrid Model for Improving Fraud Detection Accuracy in E-Commerce Financial Services. In 2025 2nd International Conference on Intelligent Algorithms for Computational Intelligence Systems (IACIS) (pp. 1-6). IEEE.
28. Moreau, L., & Groth, P. (2013). Provenance: An introduction to PROV. Morgan & Claypool.
29. Nagabhyru, K. C., Rani, M., Reddy, D. S., & Krishnaraj, V. (2025, August). Machine Learning-Driven Fault Detection in Electric Vehicles via Hybrid Reinforcement Learning Model. In 2025 2nd International Conference on Intelligent Algorithms for Computational Intelligence Systems (IACIS) (pp. 1-6). IEEE.
30. Obermeyer, Z., & Emanuel, E. (2016). Predicting the future—Big data and clinical medicine. NEJM, 375, 1216–1219.
31. Vijaya Rama Raju Gottimukkala. (2025). Agentic AI for Next-Generation Cross-Border Payments: Contextual Learning in Transaction Routing. Journal of Informatics Education and Research, 5(4). Retrieved from https://jier.org/index.php/journal/article/view/3794
32. Pearl, J. (2009). Causality (2nd ed.). Cambridge University Press.
33. Srikanth, T., Segireddy, A. R., & Elavarasi, S. A. (2025, October). STaSFormer-SGAD: Semantic Triplet-Aware Spatial Flow-Guided Spatio-Temporal Graph for Anomaly Detection in Surveillance Videos. In 2025 International Conference on Communication, Computer, and Information Technology (IC3IT) (pp. 1-7). IEEE.
34. Rajkomar, A., Dean, J., & Kohane, I. (2019). Machine learning in medicine. NEJM, 380, 1347–1358.
35. Amistapuram, K. (2025). Agentic AI for Next-Generation Insurance Platforms: Autonomous Decision-Making in Claims and Policy Servicing. Journal of Marketing & Social Research, 2, 88-103.

ISSN: 3067-4166

AMERICAN DATA SCIENCE JOURNAL FOR ADVANCED COMPUTATIONS

VOLUME: 03 ISSUE: 03

RECEIVED: JULY 08 REVISED: AUGUST 04 ACCEPTED: AUGUST 24 PUBLISHED: SEPTEMBER 12

36. Rieke, N., Hancox, J., Li, W., et al. (2020). Federated learning for digital health. NPJ Digital Medicine, 3, 119.
37. Annapareddy, V. N., Singireddy, J., Preethish Nanan, B., & Burugulla, J. K. R. (2025). Emotional Intelligence in Artificial Agents: Leveraging Deep Multimodal Big Data for Contextual Social Interaction and Adaptive Behavioral Modelling. Jai Kiran Reddy, Emotional Intelligence in Artificial Agents: Leveraging Deep Multimodal Big Data for Contextual Social Interaction and Adaptive Behavioral Modelling (April 14, 2025).
38. Amershi, S., Begel, A., Bird, C., et al. (2019). Software engineering for machine learning: A case study. Proceedings of the International Conference on Software Engineering, 291–300.
39. Pamisetty, A., Paleti, S., Adusupalli, B., Singireddy, J., Inala, R., & Nagabhyru, K. C. (2025, September). Explainable AI Systems for Credit Scoring and Loan Risk Assessment in Digital Banking Platforms. In 2025 IEEE 13th International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications (IDAACS) (pp. 1478-1483). IEEE.
40. Armbrust, M., Zaharia, M., Xin, R. S., et al. (2015). Apache Spark: A unified engine for big data processing. Communications of the ACM, 59(11), 56–65.
41. Garapati, R. S. (2025). An Intelligent IoT Security System: Cloud-Native Architecture with Real-Time AI Threat Detection and Web Visualization. Journal homepage: https://jmsronline. com, 2(06).
42. Challa, K. (2025). Innovations in Digital Finance and Intelligent Technologies: A Deep Dive into AI, Machine Learning, Cloud Computing, and Big Data in Transforming Global Payments and Financial Services. Deep Science Publishing.
43. Babaiah, C., Dobriyal, N., Shamila, M., Aitha, A. R., Patel, S. P., & Upodhyay, D. (2025, December). Intelligent Fault Detection and Recovery in Wireless Sensor Networks Using AI. In 2025 IEEE 5th International Conference on ICT in Business Industry & Government (ICTBIG) (pp. 1-6). IEEE.
44. Benjamens, S., Dhunnoo, P., & Meskó, B. (2020). The state of artificial intelligence-based FDA-approved medical devices. NPJ Digital Medicine, 3, 118.
45. Nagabhyru, K. C. (2025). Beyond Automation: The 2025 Role of Agentic AI in Autonomous Data Engineering and Adaptive Enterprise Systems.
46. Bertsekas, D. P. (2012). Dynamic programming and optimal control (Vol. 1). Athena Scientific.
47. Vajpayee, A., Khan, S., Gottimukkala, V. R. R., Sharma, D., & Seshasai, S. J. (2025). Digital Financial Literacy 4.0: Consumer Readiness for AI-Driven Fintech and Blockchain Ecosystems. International Insurance Law Review, 33(S5), 963-973.
48. Nandan, B. P. (2022). AI-Powered Fault Detection In Semiconductor Fabrication: A Data-Centric Perspective.Lebcir, I., Mageswari, S. U., Bhosale, Y. H., Nagubandi, A. R., & Mahabooba, M. M. Agile Strategic Management in the Age of Disruption: Leveraging AI and Data Analytics for Competitive Advantage.

ISSN: 3067-4166

AMERICAN DATA SCIENCE JOURNAL FOR ADVANCED COMPUTATIONS

VOLUME: 03 ISSUE: 03

RECEIVED: JULY 08 REVISED: AUGUST 04 ACCEPTED: AUGUST 24 PUBLISHED: SEPTEMBER 12

49. Satyanarayanan, M. (2017). The emergence of edge computing. Computer, 50(1), 30–39.
50. Yandamuri, U. S. (2023). An Intelligent Analytics Framework Combining Big Data and Machine Learning for Business Forecasting. International Journal Of Finance, 36(6), 682-706.
51. Sheller, M. J., Reina, G. A., Edwards, B., et al. (2020). Multi-institutional deep learning without sharing patient data. Brainlesion Workshop.
52. GUNTUPALLI, R. (2025). EXPLAINABLE AI IN CLINICAL DECISION SUPPORT: INTERPRETABLE NEURAL MODELS FOR TRUSTWORTHY HEALTHCARE AUTOMATIONEXPLAINABLE AI IN CLINICAL DECISION SUPPORT: INTERPRETABLE NEURAL MODELS FOR TRUSTWORTHY HEALTHCARE AUTOMATION. TPM–Testing, Psychometrics, Methodology in Applied Psychology, 32(S9 (2025): Posted 15 December), 462-471.
53. Shortliffe, E. H., & Sepúlveda, M. J. (2018). Clinical decision support in the era of AI. JAMA, 320(21), 2199–2200.
54. Challa, K., Sriram, H. K., & Gadi, A. L. (2025). Leveraging AI, ML, and Gen AI in Automotive and Financial Services: Data-Driven Approaches to Insurance, Payments, Identity Protection, and Sustainable Innovation.
55. Sutton, R. S., & Barto, A. G. (2018). Reinforcement learning (2nd ed.). MIT Press.
56. Siva Hemanth Kolla. (2023). Deep Learning–Driven Retrieval-Augmented Generation for Enterprise ITSM Automation: A Governance-Aligned Large Language Model Architecture . Journal of Computational Analysis and Applications (JoCAAA), 31(4), 2489–2502. Retrieved from https://www.eudoxuspress.com/index.php/pub/article/view/4774
57. Tsamados, A., Aggarwal, N., Cowls, J., et al. (2022). The ethics of algorithms. AI & Society, 37, 215–230.
58. Davuluri, P. S. L. N. . (2024). AI-Driven Data Governance Frameworks for Automated Regulatory Reporting and Audit Readiness. Metallurgical and Materials Engineering, 30(4), 996–1010. Retrieved from https://metall-mater-eng.com/index.php/home/article/view/1936
59. Wooldridge, M. (2009). An introduction to multiagent systems (2nd ed.). Wiley.
60. Bandi, V. D. V. K. (2023). Production-Grade Machine Learning Pipelines For Healthcare Predictive Analytics. South Eastern European Journal of Public Health, 189–205. Retrieved from https://www.seejph.com/index.php/seejph/article/view/7057
61. Zhang, A., Xing, L., Zou, J., & Wu, J. C. (2022). Shifting ML for healthcare to deployment. Nature Biomedical Engineering, 6, 1330–1345.
62. Kolla, S. K. (2021). Architectural Frameworks for Large-Scale Electronic Health Record Data Platforms. Current Research in Public Health, 1(1), 1–19. Retrieved from https://www.scipublications.com/journal/index.php/crph/article/view/1372
63. Benford, S., et al. (2009). Emergent multi-agent architectures. Autonomous Agents and Multi-Agent Systems, 18, 15–45.

ISSN: 3067-4166

AMERICAN DATA SCIENCE JOURNAL FOR ADVANCED COMPUTATIONS

VOLUME: 03 ISSUE: 03

RECEIVED: JULY 08 REVISED: AUGUST 04 ACCEPTED: AUGUST 24 PUBLISHED: SEPTEMBER 12

64. Guntupalli, R. (2025). Multi-Cloud vs. Hybrid Cloud Security: Key Challenges and Best Practices. Hybrid Cloud Security: Key Challenges and Best Practices (November 21, 2025).
65. Mell, P., & Grance, T. (2011). The NIST definition of cloud computing. NIST.
66. Kumar, S. S., Singireddy, S., Nanan, B. P., Recharla, M., Gadi, A. L., & Paleti, S. (2025). Optimizing edge computing for big data processing in smart cities. Metallurgical and Materials Engineering, 31(3), 31-39.
67. World Health Organization. (2021). Ethics and governance of artificial intelligence for health. WHO Press.
68. Kolla, S. H. (2024). RETRIEVAL-AUGMENTED GENERATION WITH SMALL LLMS FOR KNOWLEDGE-DRIVEN DECISION AUTOMATION IN ENTERPRISE SERVICE PLATFORMS. Turkish Journal of Computer and Mathematics Education (TURCOMAT), 15(3), 476–486. https://doi.org/10.61841/turcomat.v15i3.15497
69. Moreau, L., et al. (2015). The W3C PROV family of specifications. Future Generation Computer Systems, 29(7), 161–165.
70. Davuluri, P. N. Integrating Artificial Intelligence into Event-Driven Financial Crime Compliance Platforms.
71. Van Roy, P. (2009). Self-management in distributed systems. IEEE Computer, 42(12), 40–47.
72. Vardhan Kumar Bandi, V. D. (2024). Automated Feature Engineering Systems in Large-Scale Healthcare Data Environments. Journal of Neonatal Surgery, 13(1), 2127–2141. Retrieved from https://www.jneonatalsurg.com/index.php/jns/article/view/10004
73. Sutton, R. S. (2019). The bitter lesson. Incomplete Ideas Blog.
74. Singireddy, S. (2025, May). AI-Driven Comprehensive Insurance and AAA Membership Benefits Overview. In 2025 2nd International Conference on Research Methodologies in Knowledge Management, Artificial Intelligence and Telecommunication Engineering (RMKMATE) (pp. 1-13). IEEE.
75. Brundage, M., Avin, S., Clark, J., et al. (2018). The malicious use of artificial intelligence. arXiv.
76. Nigam, N., Sireesha, B., Ediga, P., Segireddy, A. R., & Bokde, S. (2025, December). Comparative Evaluation of Cloud Security Algorithms Using Multiple Classifiers with an Optimized Intrusion Detection System. In 2025 IEEE 5th International Conference on ICT in Business Industry & Government (ICTBIG) (pp. 1-6). IEEE.
77. Chen, M., Mao, S., & Liu, Y. (2014). Big data: A survey. Mobile Networks and Applications, 19, 171–209.
78. Pareyani, S., Goswami, S., Geetha, Y., Dimri, S. K., Niharika, D. S., & Amistapuram, K. (2025, December). Smart Resource Allocation in Wireless Sensor Networks Through AI Techniques. In 2025 IEEE 5th International Conference on ICT in Business Industry & Government (ICTBIG) (pp. 1-6). IEEE.
79. Dehghani, Z. (2022). Data mesh. O'Reilly Media.

ISSN: 3067-4166

AMERICAN DATA SCIENCE JOURNAL FOR ADVANCED COMPUTATIONS

VOLUME: 03 ISSUE: 03

RECEIVED: JULY 08 REVISED: AUGUST 04 ACCEPTED: AUGUST 24 PUBLISHED: SEPTEMBER 12

80. Sheelam, G. K., & Nandan, B. P. (2022). Integrating AI And Data Engineering For Intelligent Semiconductor Chip Design And Optimization. Migration Letters, 19, 2178-2207.
81. Dwork, C., & Roth, A. (2014). The algorithmic foundations of differential privacy. Foundations and Trends in Theoretical Computer Science, 9(3–4), 211–407.
82. Nagubandi, A. R. (2025). Cryptocurrency Market Spillovers: Risk Contagion Across Global Financial Systems.
83. European Parliament and Council of the European Union. (2016). General Data Protection Regulation (GDPR). Official Journal of the European Union.
84. Yandamuri, U. S. AI-Driven Decision Support Systems for Operational Optimization in Hospitality Technology.
85. Gentry, C. (2009). A fully homomorphic encryption scheme. Stanford University.
86. Guntupalli, R. (2025). Federated Deep Learning for Predictive Healthcare: A Privacy-Preserving AI Framework on Cloud-Native Infrastructure. Vascular and Endovascular Review, 8(16s), 200-210.
87. Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep learning. MIT Press.
88. Dutta, P., Mondal, A., Vadisetty, R., Polamarasetti, A., Guntupalli, R., & Rongali, S. K. (2025). A novel deep learning rule-based spike neural network (SNN) classification approach for diagnosis of intracranial tumors. International Journal of Information Technology, 17(9), 5705-5712.
89. Bach, P., Kurz, M. S., & Chernozhukov, V. (2024). Causal machine learning for supply chain risk prediction and intervention planning. International Journal of Production Research, 63(1), 12-34. https://doi.org/10.1080/00207543.2025.2458121.
90. Enterprise-Scale Gen AI Orchestration Using Small LMs and LLM Agents for Intelligent ITSM and HRSD Automation in Enterprise Ecosystems. (2025). MSW Management Journal, 35(2), 1889-1897.
91. Aljohani, A. (2023). Predictive analytics and machine learning for real-time supply chain risk mitigation and agility. Sustainability, 15(20), 15088. https://doi.org/10.3390/su152015088.
92. FinOps Strategies for AI-Enabled Real-Time Compliance Platforms in Cloud Native Environments. (2025). MSW Management Journal, 35(2), 2080-2088.
93. IBM. (2023). Data fabric architecture overview. IBM Redbooks.
94. Velangani Divya Vardhan Kumar Bandi. (2024). Intelligent Data Platforms For Personalized Retail Analytics At Scale. Metallurgical and Materials Engineering, 30(4), 1011–1027. Retrieved from https://metall-mater-eng.com/index.php/home/article/view/1011-1027
95. Annapareddy, V. N., Challa, K., Komaragiri, V. B., Sriram, H. K., Kalisetty, S., & Goma, T. (2025, April). Explaining AI Techniques such as SHAP LIME and RISE in Limited Sample Size Neuroimaging Studies. In 2025 4th OPJU International Technology

ISSN: 3067-4166

AMERICAN DATA SCIENCE JOURNAL FOR ADVANCED COMPUTATIONS

VOLUME: 03 ISSUE: 03

RECEIVED: JULY 08 REVISED: AUGUST 04 ACCEPTED: AUGUST 24 PUBLISHED: SEPTEMBER 12

Conference (OTCON) on Smart Computing for Innovation and Advancement in Industry 5.0 (pp. 1-6). IEEE.