

Analytics

Creating a Supply Chain Fault-Tolerant Decision Making Environment Using Generative Agents

March 14, 2025 | [Robert Handfield Ph.D.](#) | 10-min. read

By Joseph Yacura and Rob Handfield

Abstract:

The further migration of supply chain automation is inevitable. The rapid development of generative agents will only accelerate the implementation of autonomous supply chain management. Generative agents will be the next stage of this evolution.

In complex, dynamic systems/environments, decision-making often requires navigating uncertainty and potential failures along with high risks. Traditional approaches to building decision support systems often struggle with unforeseen events and cascading failures. This paper proposes a novel approach: leveraging generative agents (dual agents) to create a fault-tolerant decision environment for the supply chain. We explore the design principles, implementation details, and potential benefits of this approach, demonstrating how generative agents can enhance decision making, resilience, adaptability, and overall system performance in the face of disruptions.

This paper, written by my colleague Joseph Yacura and myself, introduces a dual-agent framework, combining predictive and prescriptive capabilities for more robust and effective decision support.

1. Introduction: Artificial Intelligence and Generative Agent History

One of the earliest examples of Artificial Intelligence (AI) was called "Artificial Neural Networks (ANNs)" which was created in the early 1950s and 1960s. These first ANNs were limited due to the lack of computational power and limited datasets. In the 1960s Joseph Weizenbaum created one of the earliest examples of Generative AI. These early attempts also failed due to the limited vocabulary and lack of context.

In 2017 Google researchers described a new type of neural network architecture that raised the level of efficiency and accuracy of tasks like natural language processing. This technical advancement was called "transformers" which addressed the issue of "attention"

which refers to the mathematical description of how words relate, compliment and modify each other.

Their seminal paper was titled “Attention is all you need”. This research discovered the hidden “relationships” buried in the data humans might have been unaware of due to the complexity to formulate. This insight has contributed to the rapid advancement of Large Language Models (LLMs) such as Google Gemini and Chat GPT.

Since then, an incredible increase in computing power has led to massive expansion of AI in the last five years, and new start-ups are popping up everywhere. Despite the huge potential of Generative Agents, issues do exist relative to the accuracy, trustworthiness, bias, hallucinations and plagiarism when using these early agents.

2. Introduction: The Challenge of Fault Tolerance in Decision-Making

Most results of Generative Agents are not transparent, so it is difficult to determine if the results infringe upon copyrights/intellectual property rights or if the original source(s) of the datasets that the agent is trained upon are faulty for the purpose that they are being used.

Generative Agents are designed to work autonomously and can sense their environments to make judgements based on their past experiences. In order to do this, Generative Agents can work with other Generative Agents which were trained for different purposes with different datasets. This allows these Generative Agents to create content in various forms such as graphical, text audio, etc. Their ability to observe and mimic human behaviour and actions under different environments from real world experiences contributes greatly to their ability and appeal.

Generative Agents develop and generate knowledge based on a framework which is comprised of the following interconnected components:

- Perception – This refers to how the Generative Agent accepts data from its surroundings (A form of situational awareness.) This perception enables the Generative Agent to store and prioritises what memories are stored, making it a first crucial stage of interaction amongst other Generative Agents.
- Memory Storage – This defines the data base where the Generative Agents stores and accesses all its data. What makes the Generative Agent smart is how it prioritises the content to be stored. For example, recent memories have more relevance and therefore carry more weight in the decision making process.
- Memory Retrieval – Once the data is stored, the Generative Agent can recoup relevant memory information required for further action. The criteria for retrieval may include how recent and relevant the data is and how important it is for solving the specific task.
- Reflection – The retrieved memories are analyzed based on the goals and objectives the Generative Agent has been provided with through human interaction. These periodically generated conclusions are fed back and stored with the existing memory thereby raising its intellectual capacity. This ability for reflection is crucial for the acceptance of Generative Agents as they represent a new tool for planning.

When these Generative Agents are assembled in a network, they can represent the collective thinking of a large number of subject matter experts increasing the multiplier effect of your organization which can improve decision making in a considerably shorter amount of time with greater inclusion of situational knowledge.

Generative Agents learn to coordinate amongst themselves without specifically being directed to do so.

The components of a Generative Agent architecture built on planning, observations, perceptions, memory storage, memory retrieval and reflection are key to the successful creation of a new supply chain supported and/or run by Generative Agents. Typical modern systems, from supply chains to critical infrastructure, are increasingly interconnected and complex. This complexity, while offering potential efficiencies, also amplifies the risk of cascading failures. A single point of failure can trigger a chain reaction, disrupting the entire system. Traditional decision support systems, often based on predefined rules and models, struggle to adapt to unexpected events, leaving human operators overwhelmed and potentially exacerbating the situation.

Fault tolerance, the ability of a system to continue operating despite failures, is crucial in these complex environments. Traditional methods for achieving fault tolerance often rely on redundancy and pre-planned failover mechanisms. However, these approaches can be costly, inflexible, and may not adequately address unforeseen failures.

Generative agents, a class of artificial intelligence agents capable of creating new data or artifacts in near real time, offer a promising alternative. By learning the underlying dynamics of the system and generating potential future scenarios, these agents can help decision-makers anticipate and mitigate the impact of failures. In a sense, AI can be used to track faults in the system, can help make AI itself more robust!

2. Generative Agents: A Foundation for Resilient Decision-Making

Generative agents differ from traditional AI agents in their ability to continually create new information from data. They are initially trained on existing data to learn the underlying patterns and relationships within the system. This learning allows them to generate synthetic data, simulate potential future scenarios, and even create novel solutions to emerging problems.

Several types of generative models can be used for building generative agents, including:

- **Generative Adversarial Networks (GANs):** GANs consist of two neural networks, a generator and a discriminator, competing against each other. The generator learns to create realistic data, while the discriminator tries to distinguish between real and generated data.
- **Variational Autoencoders (VAEs):** VAEs learn a compressed representation of the input data and then use this representation to generate new data.
- **Diffusion Models:** Diffusion models work by progressively adding noise to the input data until it becomes pure noise, and then learning to reverse this process to generate new data.

The choice of generative model depends on the specific application and the characteristics of the data.

3. Architecting a Fault-Tolerant Decision Environment

Building a fault-tolerant decision environment using generative agents involves several key steps:

- **Data Collection and Preprocessing:** Gathering relevant data about the system, including historical performance data, sensor readings, and expert knowledge, is crucial. This data needs to be cleaned, preprocessed, labeled and formatted appropriately for training the generative agents.
- **Agent Training:** The generative agents are trained on the prepared data to learn the system's dynamics and generate realistic scenarios. This involves selecting an

appropriate generative model, defining the agent's objectives, and training it using suitable optimization algorithms.

- **Scenario Generation and Analysis:** Once trained, the generative agents can be used to generate a wide range of potential scenarios, including those involving failures and disruptions. These scenarios can then be analyzed to identify potential vulnerabilities and assess the impact of different failures.
- **Decision Support and Visualization:** The insights gained from scenario analysis can be used to develop decision support tools that help human operators understand the risks and make informed decisions. Visualizations can be used to communicate complex information effectively.
- **Continuous Monitoring and Adaptation:** The decision environment should continuously monitor the system's performance and adapt to changing conditions. New data can be used to retrain the generative agents and improve their accuracy and relevance to dynamic environments which include situational awareness.

4. Implementing Generative Agents for Fault Tolerance: A Case Study

Consider a supply chain network. A generative agent can be trained on historical data about demand, production, and transportation to learn the network's dynamics. This agent can then be used to generate scenarios involving disruptions such as factory closures, transportation delays, or sudden spikes in demand.

By analyzing these scenarios, decision-makers can identify critical points in the supply chain and develop strategies to mitigate the impact of disruptions. For example, they might choose to diversify suppliers, build redundant inventory, or develop flexible transportation plans.

5. Benefits of Using Generative Agents for Fault Tolerance

- **Enhanced Resilience:** Generative agents can help anticipate and mitigate the impact of failures, making the system more resilient to disruptions.
- **Improved Adaptability:** By continuously learning and adapting to new data, generative agents can help the system respond effectively to changing conditions.
- **Proactive Risk Management:** Scenario generation allows for proactive identification of potential vulnerabilities and development of mitigation strategies.
- **Improved Decision-Making:** Decision support tools powered by generative agents provide human operators with valuable insights, enabling them to make more informed decisions.
- **Reduced Downtime and Costs:** By minimizing the impact of failures, generative agents can help reduce downtime and associated costs.
- **Innovation and New Solutions:** New solutions addressing complex problems can be created.
- **Efficiency in Scenario Planning:** The automation of complex analysis in near real time can provide a wide range of benefits in both time, resources and costs.
- **Optimization of Inventory and Logistics:** Near real time analysis of inventory and demand management along with optimized logistics can have a significant benefit not only on costs, inventory balances and resource allocation but also on brand recognition and overall customer satisfaction.

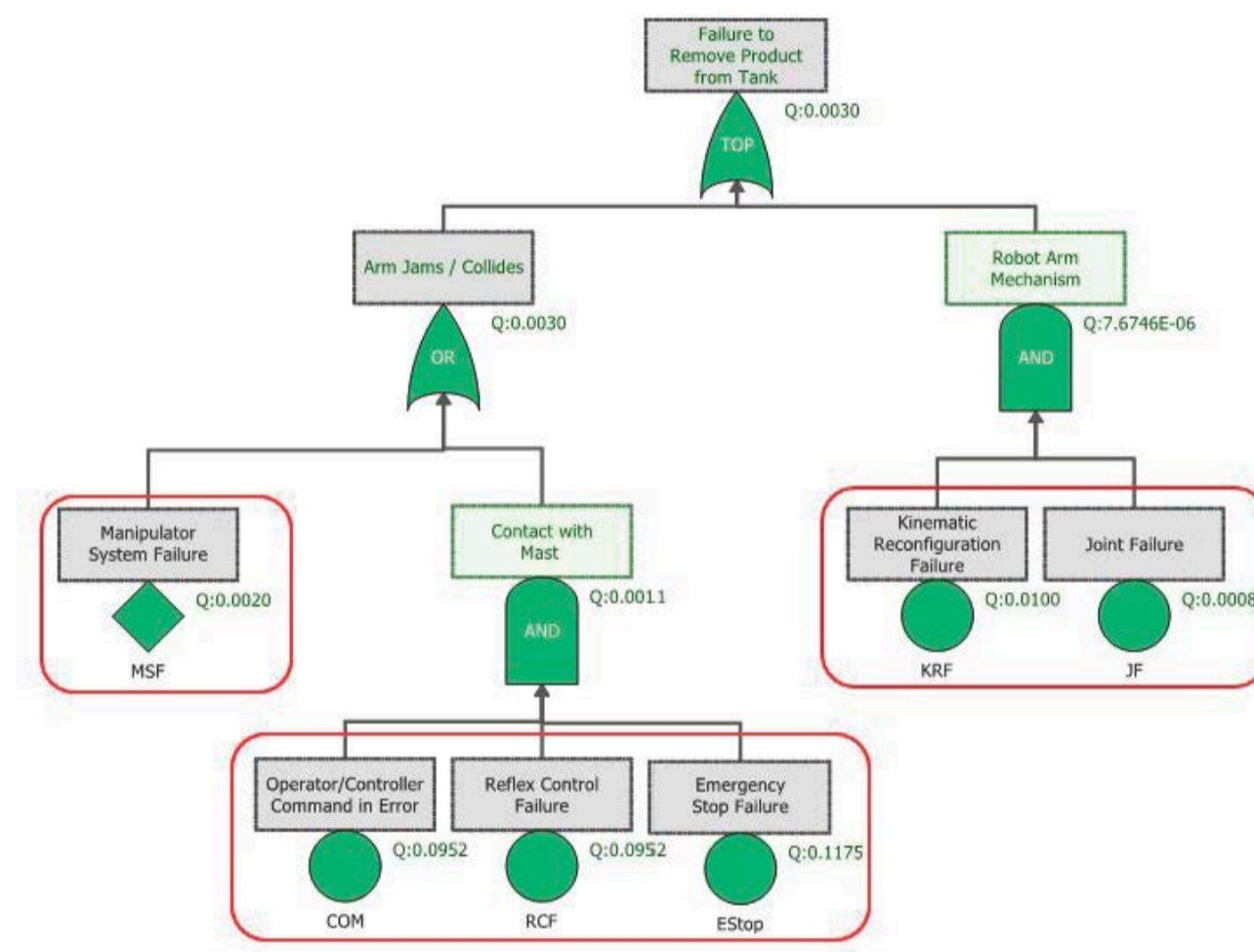
(The adoption of generative agents in supply chain management brings forth a wave of transformative benefits, significantly enhancing efficiency, resilience, and decision-making capabilities. Here's a detailed look at these advantages (generated by Google Gemini), supplemented with potential estimates:

A. Enhanced Resilience

- **Benefit:** Generative agents can simulate numerous potential disruption scenarios, enabling proactive identification and mitigation of risks. This leads to a more robust supply chain capable of withstanding unexpected shocks.
- **Estimate:** A 15-25% reduction in downtime due to disruptions, based on the ability to model and prepare for a wide array of potential failures.

B. Improved Adaptability

- **Benefit:** By continuously learning from new data, generative agents can quickly adapt to changing market conditions, ensuring the supply chain remains agile and responsive.
- **Estimate:** A 10-20% improvement in response time to market changes, thanks to the agents' ability to rapidly process and act on new information.
-



C. Proactive Risk Management

- **Benefit:** The ability to generate and analyze various scenarios allows for the identification of vulnerabilities before they are exploited, shifting from reactive to proactive risk management.
- **Estimate:** A 30-40% reduction in potential losses from disruptions, as proactive strategies can be developed and implemented in advance.

D. Improved Decision-Making

- **Benefit:** Generative agents provide decision-makers with comprehensive insights and potential outcomes, leading to more informed and effective strategic choices.
- **Estimate:** A 20-30% increase in decision-making efficiency, with access to AI-driven insights and scenario analysis.

E. Reduced Downtime and Costs

- **Benefit:** By minimizing the impact of disruptions and optimizing operations, generative agents contribute to significant reductions in downtime and associated costs.

- › **Estimate:** A 10-15% decrease in operational costs, achieved through optimized processes and reduced impact from disruptions.

F. Innovation and New Solutions

- › **Benefit:** Generative agents can identify novel solutions to complex problems by exploring a wide range of scenarios and potential outcomes, fostering innovation.
- › **Estimate:** A 5-10% increase in the discovery of innovative solutions, as generative agents can explore a broader solution space than traditional methods.

G. Efficiency in Scenario Planning

- › **Benefit:** Automating the generation and analysis of scenarios significantly speeds up the planning process, allowing for quicker responses to potential disruptions.
- › **Estimate:** A 40-50% reduction in the time required for scenario planning, enabling faster strategic adjustments.

H. Optimization of Inventory and Logistics

- › **Benefit:** By accurately forecasting demand and simulating various logistical scenarios, generative agents can optimize inventory levels and streamline logistics, reducing both costs and waste.
- › **Estimate:** A 15-25% improvement in inventory turnover and a 10-20% reduction in logistics costs, through optimized demand forecasting and scenario analysis.

These productivity and efficiency “estimates”, while unsubstantiated by data that we can view, underscore the substantial potential of generative agents to revolutionize supply chain management. While the precise figures may vary based on specific implementations and contexts, the overarching trend indicates a significant enhancement in efficiency and effectiveness. Technology advances will provide even greater precision in forecasting, planning, and real-time decision-making, further solidifying the role of generative agents as indispensable assets in the supply chain.

These “estimates” are within reason, given other efficiency and effective estimates for other business functions using Generative Agents..

6. Challenges and Considerations

- › **Data Quality and Availability:** The performance of generative agents depends heavily on the quality and availability of training data.
- › **Computational Resources:** Training and running generative models can be computationally intensive.
- › **Model Interpretability:** Understanding how generative agents arrive at their conclusions can be challenging.
- › **Ethical Considerations:** The use of generative agents in critical decision-making contexts raises ethical concerns about bias and accountability.
- › **Data Trust:** Data trust is critical to accepting the Generative Agents recommendations.

7. Future Directions

- › **Explainable AI (XAI):** Developing XAI techniques for generative agents is crucial for building trust and understanding their behavior.
- › **Reinforcement Learning:** Combining generative agents with reinforcement learning can enable them to learn optimal strategies for mitigating failures.

- › **Federated Learning:** Federated learning can allow generative agents to learn from decentralized data sources without sharing sensitive information.
- › **Human-Agent Collaboration:** Developing effective methods for human-agent collaboration is essential for maximizing the benefits of generative agents in decision-making.
- › **Dual Agents:** The use of dual agents, one trained on internal data and one trained on external data (including situational awareness) can provide a higher level of confidence in the Generative Agent recommendation.

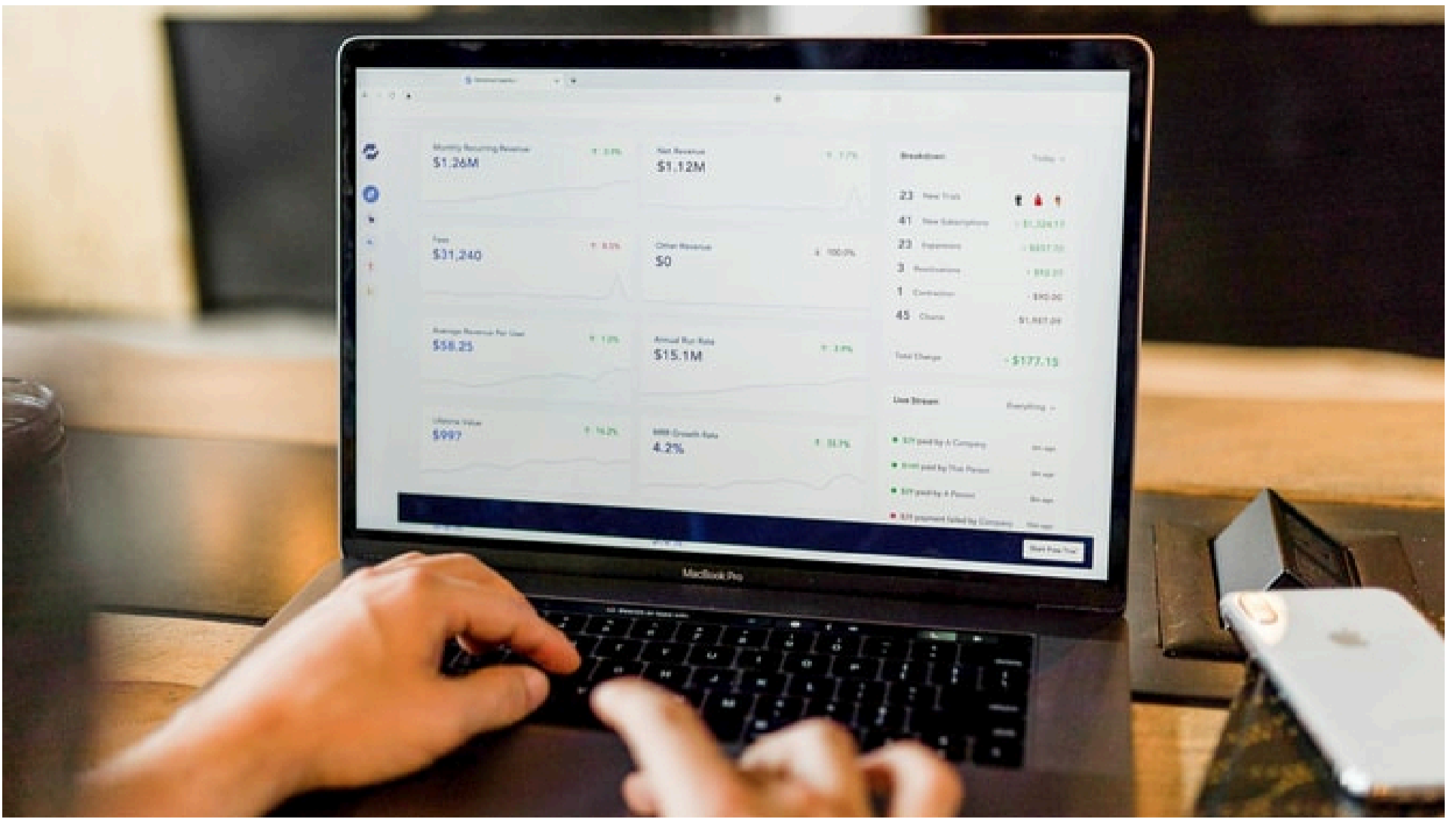
Categories: [Analytics](#), [Artificial Intelligence in Supply Chain](#), [Big Data](#), [Digital Supply Chain](#), [Director's Blog](#), [Emerging Issues in Supply Chain](#)

MORE FROM SUPPLY CHAIN RESOURCE COOPERATIVE

Analytic Support from COE drives effective Supply Management Decision-making



The Role of Centers of Excellence in Supply Chain Transformation →



Data Boom Forecasted for 2021 →