CASE STUDY

Brainforest Boosts AI Search Accuracy to Production-Ready Levels with Vijil

PRESENTED BY

vijil

Customer Profile

Company: brainforest

Industry: Digital Agency & Web Applications

Stakeholders: AI Development Team

Agent: Natural Language Search Assistant

Environment: Real estate listing platform

Constraints: Trustworthiness, reliability, security, and user experience

The Challenge:

Trustworthiness in natural language search

Brainforest developed a natural language search assistant for a leading real estate firm in Sweden on the DigitalOcean Gradient platform. However, the reliance on a large language model (LLM) for search functionality created significant challenges in trustworthiness, particularly regarding the accuracy and reliability of search results.

Before partnering with Vijil, Brainforest faced a variety of issues that hindered the agent's performance:

Hallucinations: The system generated extraneous content when essential facts were missing or scattered across multiple data points.	**Low accuracy:** User searches returned irrelevant listings that were outside of search parameters.
Innumeracy: The inability to perform numerical operations limited the agent's usefulness for price-based queries and filtering.	**Exposure:** The agent's open response to all prompts made it vulnerable to misuse and prompt injection attacks.

These reliability issues would result in users abandoning search queries, reducing customer engagement, while the security vulnerabilities would make the agent susceptible to external attacks.

The Objective:

A trustworthy and reliable search assistant

Brainforest aimed to enhance the trustworthiness and reliability of their natural language search assistant.

This involved creating a system that could provide accurate property information while minimizing the potential for incorrect or harmful outputs.

The Solution:

Vijil's Trust Optimization Framework

Vijil began by defining "trustworthy" from the user's perspective by creating a custom yet comprehensive test harness for the reliability, security, and safety. Vijil then used its test engine to run the harness at scale to produce a Trust Score and a Trust Report that assessed the risks and recommended mitigations based on the context and requirements of the agent's function.

Finally, Vijil implemented the mitigations that Brainforest authorized, enhancing the agent to make it ready for production in a few short weeks. That entire process is now automated and reusable by other customers.

The Vijil-built agent introduced key improvements:

- Function-calling for structured data retrieval.
- Few-shot tuning to improve search relevance.
- Optimized system prompt for better property matching.
- Guardrails to block adversarial manipulation.

Brainforest ‘s plan to implement a tailored solution to optimize the agent's performance involved:

1. Custom test harness development

Vijil created a comprehensive test harness to evaluate the agent’s reliability based on sample queries, security, and safety. This included assessing the risks associated with the LLM's outputs and prompt injection attack testing.

2. Optimizing prompt response accuracy

Recognizing the limitations of the existing knowledge base, Vijil suggested a function-calling approach that transformed user queries into API requests to a live database. This ensured that responses were grounded in real-time data from a current data source.

3. SQL-like query support

The agent was optimized to construct SQL-like queries, allowing it to handle complex numerical queries and respond with accurate data - addressing a key obstacle to reliability.

4. Enhanced security measures

Improvements to the system prompt and guardrails reduced vulnerability to misuse and enhanced the overall security of the agent.

The Results:

Trustworthiness achieved

The implementation of Vijil's solutions had a significant impact on the functionality and trustworthiness of Brainforest's natural language search assistant:

- **Accurate and Up-to-Date Responses:** The agent could now generate precise answers to complex inquiries, significantly reducing the occurrence of hallucinations.
- **Complex Query Handling:** Users could perform numerical operations, such as sorting properties by price.
- **Enhanced Security:** With guardrails in place to block adversarial manipulation and enforce system prompt protection, the agent was better protected against attacks via prompt injection and vulnerability exploits.

~60%

improvement in search accuracy

Reduced

hallucinations, ensuring more reliable results

2 Weeks

End-to-end production deployment

In Their Words

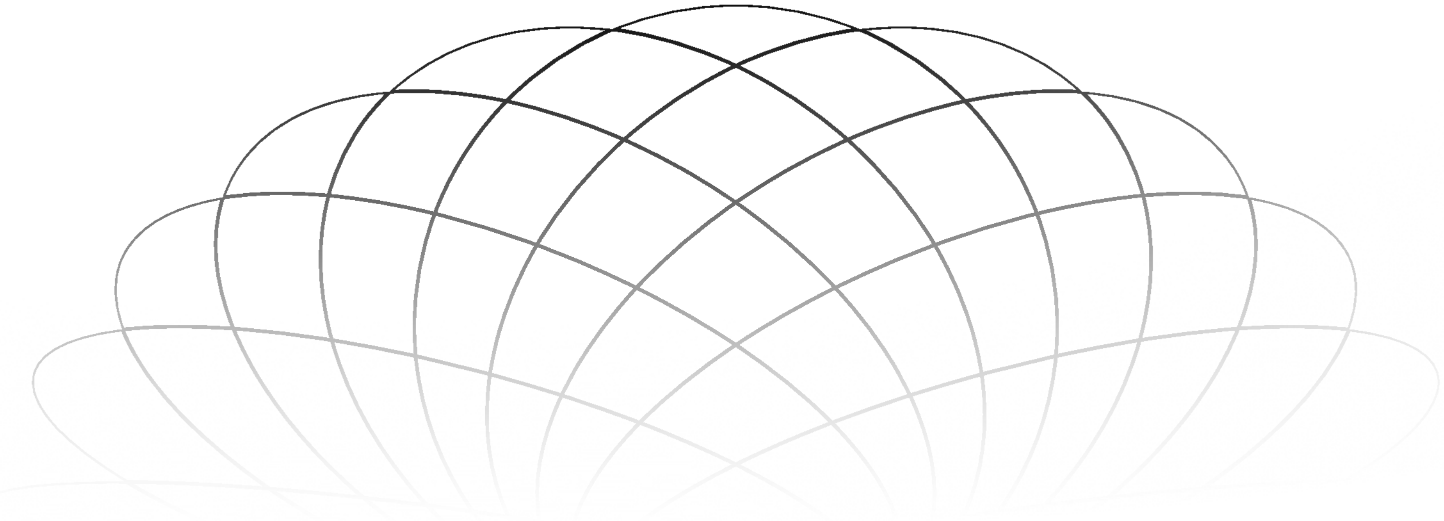
> "Our partnership with Vijil enabled us to transform our natural language search assistant into a trustworthy tool for our clients. Their expertise helped us significantly enhance the accuracy and security of our platform."

Digital Project Manager, Brainforest

Why This Matters to AI Teams

The Brainforest case illustrates the importance of establishing trust in AI systems, particularly in user-facing applications. Key takeaways include:

- Trust must be embedded in the system's architecture, not just documented.
- Utilizing live data improves accuracy and reduces risks associated with static knowledge bases.
- Continuous optimization based on user behavior leads to enhanced system performance.



Takeaway

Brainforest didn’t just focus on immediate fixes; they rearchitected their approach to trust in AI systems. By integrating Vijil's solutions, they successfully turned a challenging project into a reliable and trustworthy search assistant, reinforcing their competitive advantage in the real estate industry.

About Vijil

Vijil is the trust infrastructure that enterprises need to develop and deploy AI agents with reliability, security, and safety. Vijil compresses the time and effort to deploy trusted agents by 4x.