Implementing a Secure and Governed Real Estate AI Agent Chatbot:

A Leadership and Management Approach Using the NIST AI Risk Management Framework

Authored by Jake Kantor (2026)
jake.t.kantor@gmail.com

ABSTRACT

Real estate is a core business in the economy. Real estate will evolve with the growth of technology. According to MITRE, "Artificial intelligence (AI) technology is advancing at a rapid rate and adoption is on the rise" (MITRE Corporation, n.d.). Here in Virginia Beach, VB Real Estate, LLC is building a Real Estate AI Agent chatbot to answer listing questions for potential buyers. As the CISO, the secure implementation of this technology is of the highest importance. Karim emphasizes "the integration of AI into systems introduces numerous vulnerabilities that can jeopardize safety, security, and operational integrity" (Karim).

Considering these vulnerabilities, it is critical to implement guardrails and safeguards to protect both the real estate organization and its consumers from pain. VB Real Estate, LLC must anticipate changes, adapt security strategies, and remain resilient to properly implement AI agent chatbots into its business structure (CISSP). This is especially true when considering the implementation of AI Agents. According to Huang, "Agentic AI systems don't sit quietly in a corner—they browse, write code, spin up sub-agents, hit production APIs, and do it all at machine speed" (Huang et al.).

Traditional AppSec tools were not built for machines that reason and improvise. Therefore, the secure

architecture of the Real Estate AI Agent chatbot is crucial. (Huang et al.). Various regulatory landscapes must be considered throughout the chatbot structuring process. The Fair Housing Act must be followed by all real estate professionals and their AI outputs. NIST's AI Risk Management Framework "offers a structured process for managing organizational risks in an ever-evolving threat landscape" (Marsland) and is, therefore, a foundational component of this paper. Lastly, risk management must be taken seriously at all levels of the organization. Ideally, cultural change is achieved through senior-level organizational commitment to ensure effective risk management (NIST AI 100-1).

SYSTEM ARCHITECTURE

The Real Estate AI Agent chatbot, utilized by VB Real Estate, LLC, answers potential buyers' and sellers' questions and inquiries by pulling live listings data from the MLS. The chatbot will point users to the external showings scheduling page,

FIGURE 1 - SYSTEM PROCESS

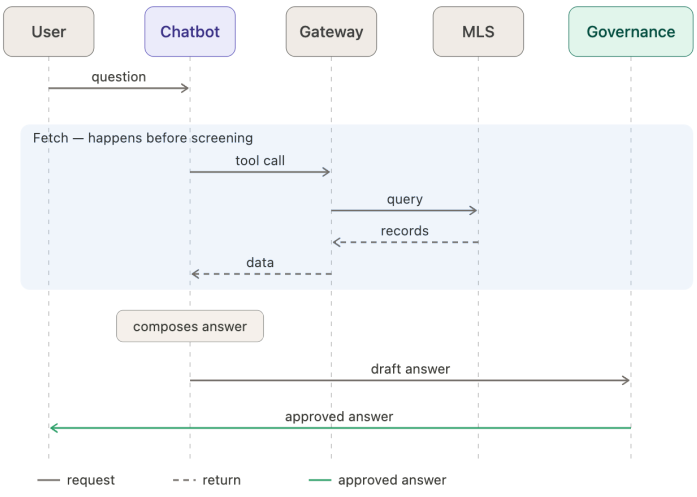
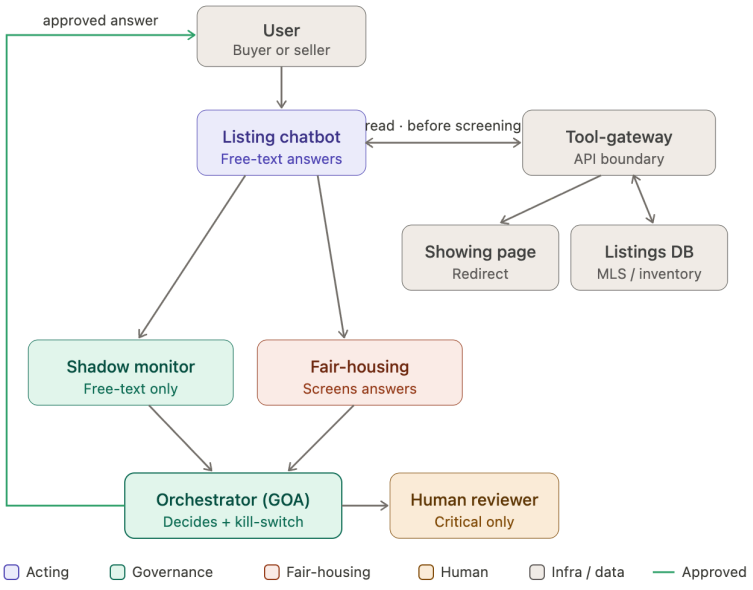


FIGURE 2 - SYSTEM ARCHITECTURE



efficiently and securely streamlining the process.

Multiple AI agents are tasked with ensuring that all outputs are compliant before reaching the user. A Fair Housing Agent screens answers for flags or wording that could potentially steer or discriminate against buyers. This check protects the human real estate agent from potential license revocation. Additionally, a Shadow Monitor Agent screens answers for vulnerabilities in the AI system, such as poisoning, hallucination, and PII data leakage.

Once the Fair Housing Agent and/or the Shadow Monitor Agent flags a chatbot response, the answer is sent to an Orchestrator Agent to decide whether to disregard the flag, terminate the chatbot response, or kill a faulty AI system. When necessary, the Orchestrator Agent will escalate the issue to a human reviewer to decide the best path forward with the problematic response. For example, the AI chatbot is not allowed to appraise the value of a home. Instead, the Agentic AI System will escalate the question to a qualified human real estate agent to assist with the appraisal.

Once all checker agents have validated the response, the chatbot produces the answer to the user. When a user requests to schedule a showing, the chatbot produces an external link via a tool-gateway to the showings or listings page. The external redirection is an intentional security decision structured to prevent agent manipulation or error.

RELATED PROJECTS

The Real Estate AI Agent chatbot architecture is inspired by both Karim's Agentic AI governance research and Huang's AAGATE framework. Karim understood that, while training Internet of Robotic Things LLM models, "privacy

breaches are a concern, as IoRT systems often process sensitive data" (Karim). Huang's AAGATE model, which uses Shadow Monitor Agents and Orchestrator Agents to screen flags in AI responses, provides a robust, end-to-end solution for potential Agentic AI drawbacks within a business (Huang et al.).

REGULATORY LANDSCAPE

Fair Housing Act

The Fair Housing Act is the bedrock of prejudice prevention amongst real estate professionals. The FHA "protects people from discrimination when they are renting or buying a home, getting a mortgage, seeking housing assistance, or engaging in other housing-related activities" (HUD). All real estate Agentic AI outputs must abide by the Fair Housing considerations. This will protect both VB Real Estate, LLC, and its consumers from pain and suffering due to noncompliant AI responses.

It is important to build a strong team to ensure AI Security and Fair Housing compliance. According to the AI RMF Playbook, "entities include diverse perspectives from technical and non-technical communities throughout the AI life cycle to anticipate and mitigate unintended consequences including potential bias and discrimination" (NIST AI RMF Playbook). Therefore, focusing on AI governance, risk, and compliance will likely ensure a compliant AI system.

Secondly, the AI chatbot must ensure fairness on a technical level. The goal of the FHA is to end illegal discrimination in nearly all forms of housing, including private, public, and federally funded (HUD). According to NIST, "Fairness in AI includes concerns for equality and equity by addressing issues such as harmful bias and discrimination" (NIST AI 100-1). The

chatbot must never discriminate nor steer a user. Building technical guardrails will be important to ensuring fairness and Fair Housing compliance within the system.

NIST AI RMF

The NIST AI Risk Management Framework provides a “voluntary, structured, and flexible foundation for addressing the multifaceted risks of AI systems" (Huang et al.). It is critical for the organization of risks associated with AI systems and the structuring of fairness and safety in AI risk management (Karim). While there have been no AI regulations passed federally within the United States, the NIST AI RMF is the standard for responsible implementation of AI. Followers of this framework will be better positioned to be compliant with future federal AI regulations. It is a “critical need for a practical platform that translates these high-level functions into concrete architectural patterns and engineering controls" (Huang et al.). There are four core functions of the NIST AI RMF: Govern, Map, Measure, and Manage.

Governance is the first step of the NIST AI RMF and is meant to inform and be infused throughout the other three functions. (NIST AI 100-1). It is the preparation stage and is essential for the strategic guidance of all AI implementations. Organizational committees are a crucial aspect of AI Governance, as “engaging stakeholders early in the process is crucial for successful RMF implementation” (Marsland). AI Governance must be auditable and traceable to its source, or it is noncompliant with NIST’s AI RMF. VB Real Estate, LLC is owned solely by the real estate agent. For this reason, he is on the governance committee and has hired one AI security expert, the CISO, to guide his vision. Once an AI system’s governance has been properly

established, with rules and strategies for future AI implementation, then the organization can move to the next step: Mapping.

Mapping determines and documents the risks of the AI system. Mapping determines the exposure level and the organization's risk reduction intentions. Much of the mapping occurs when the Fair Housing and Shadow Monitor Agents flag chatbot outputs for violations. "The Shadow Monitor Agent (SMA) acts as a continuous, internal red team, re-evaluating agent plans to detect deviations and manage risks like goal manipulation or hallucination exploitation before they cause harm… and provides a centralized point to map the system's entire interaction surface and identify cross-layer threats" (Huang et al.). Next is the Manage step in the NIST AI RMF.

The NIST Measure function includes "tracking metrics for trustworthy characteristics, social impact, and human-AI configurations. (NIST AI 100-1). Rigorous software configuration must take place to identify the "how" behind AI system risks and vulnerabilities. This step is where an AI Security professional may be most valuable. An AI governance expert who focuses on legality often does not understand how to get under the hood of these systems to prevent technical compliance violations. The software can be adjusted to prevent Fair Housing Act noncompliance and protect both VB Real Estate, LLC's license and the consumer from AI prejudice. The final step of NIST's AI RMF is Manage.

The Manage function, according to NIST, "allocates risk resources to mapped and measured risks on a regular basis" (NIST AI 100-1). This function is largely covered by the Orchestration Agent, which decides how to move forward with flagged chatbot outputs. Additionally, the option to escalate risk assessments to a human is

a direct management function. According to Karim, "the NIST AI RMF itself is more of a governance framework and thus lacks technical details about threats or mitigations" (Karim). Therefore, it is important that an AI Security expert properly manages mapped and measured risks to prevent reoccurrences in the future.

RISK MITIGATION

Harmful Bias

Harmful bias is a risk that must be assessed following the NIST AI RMF before launching a real estate AI Agent chatbot. Regarding bias, "AI systems can potentially increase the speed and scale of biases and perpetuate and amplify harms to individuals, groups, communities, organizations, and society" (NIST AI 100-1). For this reason, a strong and diverse governance team must be in place to make proper and well-thought-out decisions when mapping and managing AI risks. Harmful bias will be mapped by the Shadow Monitor and Fair Housing AI Agents, flagging potential violations for the Orchestrator Agent to decide upon. In order to measure harmful bias, an AI Security expert will need to test for fairness in the AI system. Adjustments should be made to the model based on fairness testing results. Lastly, harmful bias will be managed in real time by the Orchestrator Agent and the human-in-the-loop. The governance team will also decide upon the best path forward in order to manage a risk. This is a critical process to reduce harmful bias, because AI real estate chatbots “should be evaluated to ensure it does not produce biased recommendations or engage in unlawful steering based on protected characteristics” (HUD).

AI Hallucination

Additionally, AI hallucination poses a serious risk to any consumer-facing AI system.

"Hallucination attacks involve adversaries introducing fabricated information into the AI system, causing it to generate false outputs" (Karim). VB Real Estate, LLC's AI governing body must establish policies defining acceptable levels of hallucination risks. In order to properly map this risk, the Shadow Monitor Agent "identifies AI system limitations, detecting and tracking negative impacts and incidents, and sharing information about these issues with appropriate AI actors" (AI RMF Playbook). Proper measurements will test the accuracy of the system and the RAG retrieval rate of the Agentic AI system. Lastly, the measurements found must be adjusted to manage the risk of hallucination.

Information Security

Information security is important for any AI system. It ensures that "confidentiality, integrity, and availability through protection mechanisms that prevent unauthorized access and use may be said to be secure" (NIST AI 100-1). It is important to strategize for the future during the governance stage of information security. The Real Estate AI Agent chatbot is secured because we've implemented checker AI agents to ensure the integrity of the system. For example, "policies enforced by the Orchestration Agent explicitly forbid revealing PII" (Huang et al.). During the mapping phase, we've ensured that "no agent can act or communicate unless its identity, origin, and network path are explicitly verified" (Huang et al.). Information security is measured through continuous vulnerability and penetration testing. Upon the mapping and measurement of an information security risk, the management team must decipher the best path to reducing and even eliminating information security vulnerabilities.

Data Privacy

Data Privacy is incredibly important to protect in any system.

However, it is particularly vulnerable in a self-managed AI system. In the governance phase, “privacy values such as anonymity, confidentiality, and control generally should guide choices for AI system design, development, and deployment" (NIST AI 100-1). It is critical to have a strong AI committee, with members who can understand and assist with “data privacy laws, intellectual property rights, and ethical implications related to the system’s function and data handling" (Marsland). Upon the mapping function, the Shadow Monitor Agent’s role is to ensure that no attacker can recover sensitive user or training data. Measurement is necessary to ensure data privacy is protected. The committee must continuously monitor for personally identifiable information (PII) leakage. The management team may review how the organization has “documented the AI system's data provenance, including sources, origins, transformations, augmentations, labels, dependencies, constraints, and metadata" (AI RMF Playbook).

Human Oversight

Human oversight is the most important security implementation any AI system can make, especially at this early stage of the technology. The governance team must implement "policies and procedures… to define and differentiate roles and responsibilities for human-AI configurations and oversight of AI systems and establish policies and procedures regarding AI actor roles, and responsibilities for human oversight of deployed systems” (AI RMF Playbook). In our mapping phase, inputs and outputs are tracked and documented, allowing for human review for an emergency or a standard audit. Human oversight can be measured by tracking how often a red flag was escalated and how many false positives and negatives occurred. Human oversight is managed by ensuring that "critical decisions

trigger human review" (Huang et al.). This would include niche real estate questions that only a licensed real estate professional is qualified for.

Intellectual Property

Intellectual Property Considerations are a hot topic in AI, as humans are becoming increasingly concerned with AI infringement. The governance team must ensure policies are in place to prevent infringement of a third party's intellectual property. Additionally, the system's "training data may also be subject to copyright and should follow applicable intellectual property rights laws" (NIST AI 100-1). In the Mapping function, our Shadow Monitor Agent flags inward and outward intellectual property infringement. This can be measured through continuous testing of the model, including the "reconstruction sensitive training data" (Karim). Lastly, intellectual property infringement is reduced in our model by accessing data available on the MLS.

Residual Gaps

The final risk to be considered in this paper is residual gaps and limitations. The governing committee must acknowledge that risks cannot be completely terminated. We must continuously plan to observe and test for vulnerabilities. We must map risks after controls have been implemented and measure them. We must understand how the risks work and can be reduced. "Issues such as concept drift, AI bias and discrimination, shortcut learning or underspecification are difficult to identify using current standard AI testing processes… highly secure but unfair systems, accurate but opaque and uninterpretable systems, and inaccurate but secure, privacy-enhanced, and transparent systems are all undesirable" (AI RMF Playbook). The management team must work to improve the system while accepting that not all risks can be eliminated.

The leadership and governance strategy will be crucial for the proper implementation of this new Real Estate AI Agent technology. According to NIST, the Chief Information Security Officer, Jake Kantor, is responsible for decisions regarding AI risks and for setting the tone for security culture (NIST AI 100-1). The CISO must evaluate quantitative and qualitative security risks in the AI system and plan for all individual outcomes. The NIST AI Risk Management Framework is an excellent guide to ensure all risks have been evaluated proactively to identify and mitigate potential harms.

The CISO must lead the AI governance committee. It is crucial to hire a diverse team, both in expertise and life experience. Fairness testing will be more well-rounded when evaluated by a team from various walks of life. Building a robust security team is necessary to interpret, apply, and manage the framework's intricacies tailored to the organization's unique security requirements" (Marsland).

BUSINESS CONTINUITY PLANNING

The CISO must build a business continuity plan by identifying priorities and risks. Not all risks are made the same, and some will need to be prioritized over others. When building an impact assessment, the CISO must put a dollar value on each risk by assessing how much the vulnerability would cost the company if hacked and how often it is expected to be exposed. It is equally important, yet more difficult to put a dollar value on qualitative vulnerabilities, such as: “loss of goodwill among your client base, loss of employees to other jobs after prolonged downtime, social/ethical responsibilities to the community, and negative publicity” (CISSP). The largest threat to the Real Estate AI Agent chatbot is qualitative and difficult to quantify. However, if the

chatbot were to discriminate against a consumer or violate the Fair Housing Act in any way, the real estate agent would be liable to have their license revoked. This devastating vulnerability could single-handedly destroy the entire business. This is why it is necessary to "regularly track and monitor negative risks and benefits throughout the AI system lifecycle including in post-deployment monitoring" (AI RMF Playbook). Building a business continuity plan is grueling, but instrumental to any organization's security policy.

FUTURE SCALING

This AI chatbot has the potential to continue scaling. For example, a new AI bot could write offers and contracts for clients based on their needs and desires. This would still require checker AI bots, such as the Fair Housing and Shadow Monitor Agents, and a decider agent, such as the Orchestrator Agent, to ensure that the new Agentic AI system remains compliant. A licensed real estate human-in-the-loop must be required to ensure that the contract has been written satisfactorily and is ready to be sent.

Blockchain Implementation

Additionally, it is possible to implement the blockchain into the documentation process of the Agentic AI system to ensure optimal transparency and accountability. All four functions of the NIST AI RMF (Govern, Map, Measure, and Manage) could be documented on a public or private ledger. In an audit, this would provide an invaluable "tamper-proof record of agent identity and status" (Huang et al.).

CONCLUSION

In conclusion, it is imperative to understand the security risks and

vulnerabilities of any AI system prior to implementation. As CISO, I have provided architectural diagrams and satisfied NIST’s AI Risk Management Framework. I discussed the Business Impact Assessment and how catastrophic it would be for a real estate agent to lose their license due to an unsecured AI model. Huang says, "governing autonomous agents requires moving beyond static security checklists to a dynamic, resilient, and observable control plane." (Huang et al.), while Marsland insists, "understanding and applying these principles will lead to a more structured and effective approach to managing cybersecurity risks" (Marsland). As the CISO, I have ensured that VB Real Estate, LLC has followed these security principles and is properly protected from exposure.

References

Chapple, M., Stewart, J. M., & Gibson, D. (2018). *(ISC)² CISSP Certified Information Systems Security Professional Official Study Guide* (8th ed.). Sybex.

Huang, K., Lambros, K. ("Rock"), Huang, J., Mehmood, Y., Atta, H., Beck, J., Narajala, V. S., Baig, M. Z., Haq, M. A. U., Shahzad, N., & Gupta, B. (2025). *AAGATE: A NIST AI RMF-aligned governance platform for agentic AI* (arXiv:2510.25863v2). arXiv. https://arxiv.org/abs/2510.25863

Karim, H., Gupta, D., & Sitharaman, S. (2025). *Securing LLM workloads with NIST AI RMF in the Internet of Robotic Things*. *IEEE Access, 13*, 69631–69649. https://doi.org/10.1109/ACCESS.2025.3561235

Marsland, T. (2025). *Unveiling the NIST Risk Management Framework (RMF)*. Packt Publishing.

MITRE Corporation. (n.d.). *AI Security 101*. MITRE ATLAS. https://atlas.mitre.org/resources/ai-security-101

National Institute of Standards and Technology. (2023). *Artificial intelligence risk management framework (AI RMF 1.0)* (NIST AI 100-1). U.S. Department of Commerce. https://nvlpubs.nist.gov/nistpubs/ai/NIST.AI.100-1.pdf

National Institute of Standards and Technology. (n.d.). *NIST AI RMF playbook*. NIST AI Resource Center. https://airc.nist.gov/airmf-resources/playbook/

U.S. Department of Housing and Urban Development. (n.d.). *Housing discrimination under the Fair Housing Act*. https://www.hud.gov/helping-americans/fair-housing-act-overview