When Your Virtual Agent has got to be right

kama.ai: Trust. Empathy. Accuracy.

kama.ai

Hybrid AI Virtual Agents

The New Architecture for Responsible AI
and Brand-Safe AI-RAG Systems

https://kama.ai

Executive Summary

AI is advancing quickly. Yet it is also becoming a growing liability. A 2024 Oxford study found that ChatGPT 4 hallucinated in 58% of cases.[1] The Journal of Medical Internet Research uncovered fabricated citations in nearly 29% of responses.[2] These are not edge cases, they are structural weaknesses of probabilistic models. In sectors where trust is mandatory, one wrong answer can trigger legal risk, reputational harm, or customer loss.

Trustworthy Hybrid AI addresses this by combining deterministic Knowledge Graphs with controlled Generative AI Retrieval Augmented Generation (RAG) solutions - anchored with Trusted Collections of source data. For least risk, every responses can be governed in advance, not corrected after the fact. Human oversight ensures that brand-safe answers reach customers or employees. Where higher risk is acceptable, GenAI can be used with strict containment. And it is always done with the customer or end user made aware of the potential risks.

A risk of even a 2% error rate can be massive at scale. As an example, an enterprise can have as many as 73,000 flawed responses per year in a 10,000-query per day chatbot. With a Hybrid AI RAG solution, this exposure can be all but eliminated. Information and processes your experts know, are made available through trusted autonomous AI Agents. It locks in brand alignment and enforces trusted knowledge and process orchestration at every interaction.

What follows introduces an Hybrid AI Agent framework. It is a governed platform that balances agility and accuracy. The

model steps through features like RAG Draft Assist, GenAI's Sober Second Mind®, and adjustable risk profiles. It's not just safer - it's smarter. When trust is the differentiator, a safe Hybrid AI Agent gives enterprises full control over what their virtual agents state on their behalf.

Enterprise Challenges

58%

USING OPENAI'S CHATGPT 4 AND OTHER PUBLIC MODELS, WE SHOW THAT LLMS HALLUCINATE AT LEAST 58% OF THE TIME, STUGGLE TO PREDICT THEIR OWN HALLUCINATIONS, AND OFTEN UNCRITICALLY ACCEPT USERS INCORRECT ASSUMPTIONS.[1]

OXFORD UNIVERSITY

Large Language Models (LLMs) generate fluent text, but not verified facts. Their strength lies in automated creativity, not precision. Although Retrieval-Augmented Generation (RAG) improves data access, it does not ensure accuracy. Most LLM-RAG setups can also hallucinate or deliver incomplete context. Equally important, their responses vary over time. This makes it challenging to have 'voices' that are qualified to represent an enterprise or community.

Where it becomes dangerous is in live environments. Customer-facing agents need accurate, aligned, traceable answers. Internal tools need to support audit-readiness and legal defensibility. When the AI guesses, trust breaks. One seriously wrong or biased answer can ruin years of hard-earned brand equity.

Responsible enterprise applications can't rely on probabilistic storytelling. They need AI that speaks like their most accurate domain experts. Enterprises need AI solutions to fill this gap. When customers ask similar questions repeatedly, you need to provide the same answer. This is where a new option bringing deterministic and creative AI together, makes sense.

Enter Hybrid AI

63%

OF CONSUMERS ARE CONCERNED ABOUT POTENTIAL BIAS AND DISCRIMINATION IN AI ALGORITHMS AND DECISION-MAKING. [3]

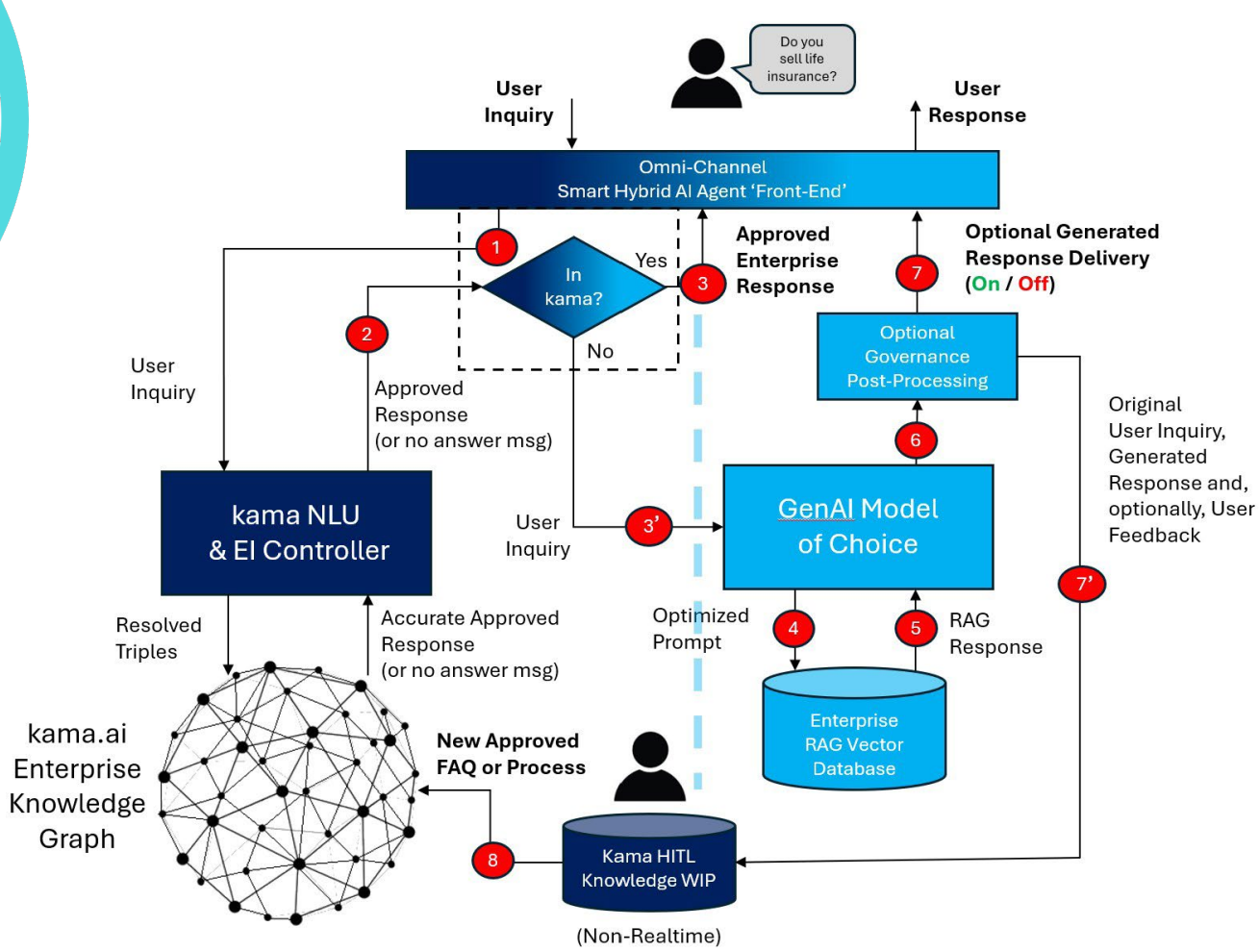
ZENDESK

Hybrid AI RAG is the core of the trustworthy AI platform discussed here. It blends deterministic Knowledge Graph tech.with tightly governed Generative AI, delivering safety and agility. The result is a smart system that answers with confidence, but only when it should.

When a user submits a question, the system checks the Enterprise Knowledge Graph first. If a verified answer exists, it's returned immediately through the front-end chatbot or medium. This ensures 100% accuracy. It is the deterministic side of the equation. Notably, it is also the safest, fastest path, and is ideal for high-trust, brand-sensitive cases.

If no answer is found, the system states this directly. The query can be routed to the generative side, if activated. If so, an answer is generated, based on the sanctioned data repository. Here, the system uses an LLM RAG chosen by the client (OpenAI, Anthropic, Gemini, Grok...). Instead of pulling from the open web or arbitrary enterprise documents, the AI exclusively focuses on the curated, Trusted Collections. These documents are hand-selected by Knowledge Administrators (KA), stored securely in enterprise repositories, and vectorized for quick retrieval without training the broader LLM model with your information.

The GenAI model produces a draft response. Then, the system can be set up in one of two ways. It might only allow the Knowledge Managers (KM) to review the draft and add it to the Enterprise Knowledge Graph of deterministic information. This means, the next time a visitor asks a similar question, the system provides the same sanctioned answer. But, the Hybrid AI Agent can also be set up such that if the answer does NOT exist, the end user will receive the generated answer based on the Trusted Collections. When used, this



second approach gives end users a clear disclaimer. It note the answer may include inaccuracies and should be verified with document links provided.

Every step is logged. Users can rate the response and provide feedback. Knowledge Managers and Administrators review flagged interactions in a periodic feedback loop. They can then review and improve collections and add approved answers into the deterministic graph. In this way the cycle reinforces itself, and the knowledge network of the organization grows over time.

This architecture, illustrated in the diagram above, turns AI from an unpredictable black box into a transparent, auditable, continuously improving knowledge and process automation system. It's not a fallback. It's a framework engineered for enterprise efficiency, intelligence, and trust.

The Cost of 2% Inaccuracy

29 %

HALLUCINATION RATES WERE, RESPECTIVELY, 39.6% (55/139), 28.6% (34/119), AND 91.4% (95/104) FOR GPT-3.5, GPT-4, AND BARD (P<.001)[2]

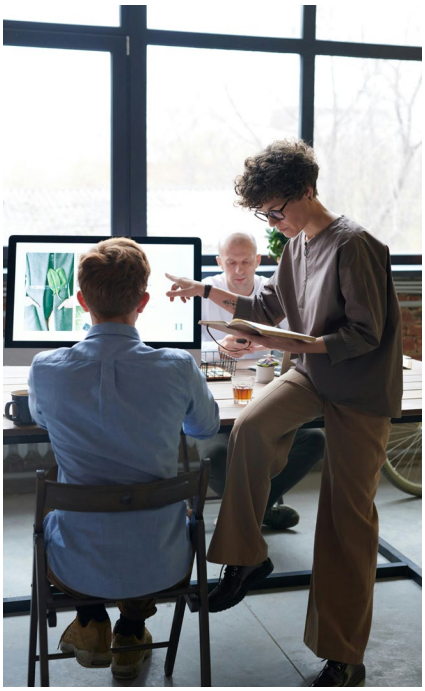
JOURNAL OF MEDICAL INTERNET RESEARCH

With AI, accuracy is not a luxury. It is a business requirement. At first glance, an AI solution that is 98% accurate might sound quite good. But in business, even a 2% error rate can be catastrophic when scaled. We hinted at this in the introduction: imagine a virtual agent handling 10,000 inquiries per day for customer service (not unusual for enterprise sized chatbots). If 2% of those interactions were answered incorrectly, that's 200 incorrect interactions every single day. These could be instances of the AI providing the wrong answer, misunderstanding the question, hallucinating, or worse offering a biased or disrespectful response.

Now multiply the 200 daily errors by 365 days. We are looking at 73,000+ failed customer interactions every year. That's 73,000 potential complaints, or rework tasks for human staff. Or perhaps these are customers who didn't get the help they needed. At volume, 2% doesn't seem so small.

Certainly, this erodes the brand confidence customers hold for the organization. Worse yet, a number of these customers will become irate. Conservatively, imagine that three in 10,000 customers take issue, and follow up with legal counsel. Between settlements and defending some lawsuits, this **can cost a typical enterprise an estimated $4.6Million annually**. Not insignificant.

This error becomes even more concerning when you consider cascading effects. Think of an AI Agents performing multiple steps in a client interaction. If there is an accuracy rate in each step of a process, it gets progressively worse with the growing number of steps in the client's case. If the accuracy of the AI agent is 90%, then



the accuracy of a two-step process is 90% x 90% or 81%. With a 3-step process, the accuracy drops to 73% and so on. Many client interactions have 5 to 10 steps. Now the chance a client gets an accurate answer drops well below 50%.

Consider that Oxford University research, again. If LLMs provide fully correct answers 42% of the time (58% of the time there are errors), we have an even more concerning issue. In this case, the probability of success of a two-step client interaction is 42% x 42% resulting in a success probability of 18%. This is truly frightful for an enterprise marketer or client service manager concerned with brand protection.

Brands cannot risk their reputations on such fragile probabilities. Especially when trust is their core currency.

Trustworthy Hybrid AI Agents reduce risk by checking answers before they reach customers. Human experts review and approve each answer and interaction. Knowledge Graphs ensure 100% verified responses. Tasks can also be automated using Robotic Process Automation. This keeps the risk levels low and gives users helpful, trusted answers and smooth experiences.

"The world does not need faster AI. It needs AI that is smarter, safer, and trustworthy."

Brian Ritchie, CEO, kama.ai

Hybrid Virtual Agent

42%

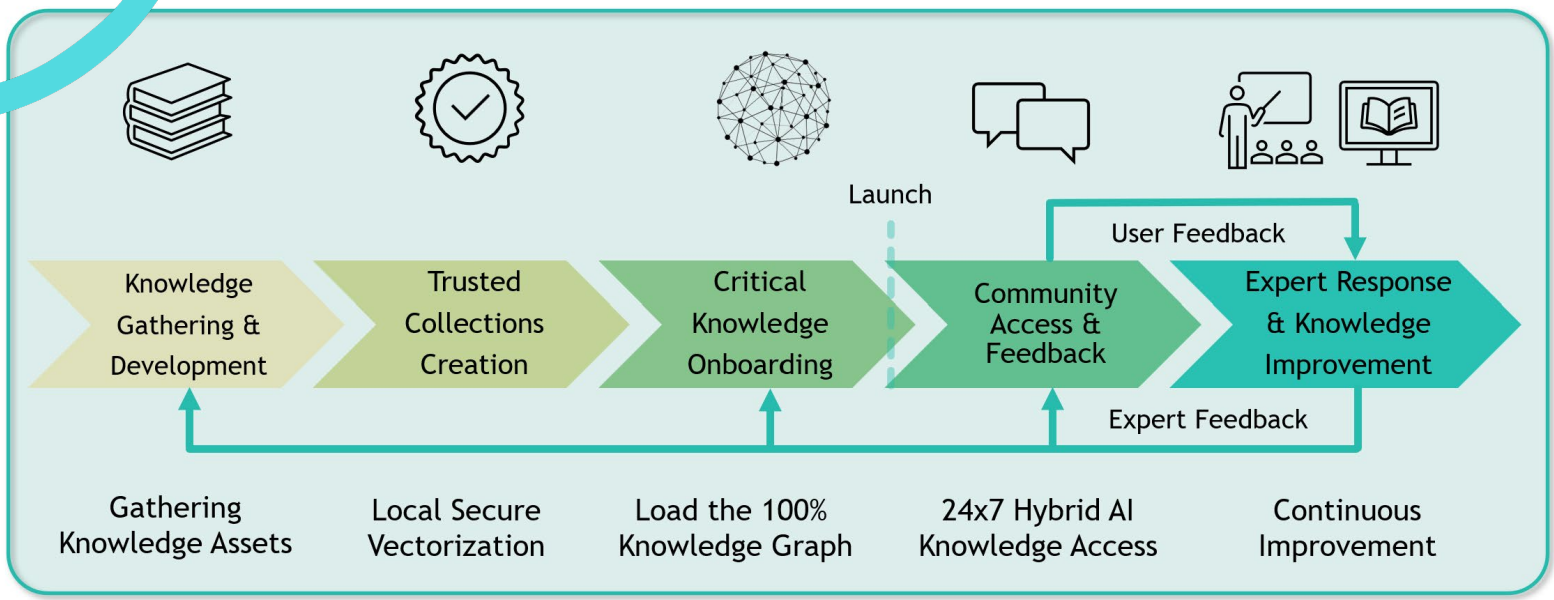
FORTY-TWO PERCENT OF CUSTOMERS TRUST BUSINESSES TO USE AI ETHICALLY — DOWN FROM 58% IN 2023. IN THIS MOMENT OF RAPID CHANGE, COMPANIES HAVE AN ESSENTIAL OPPORTUNITY TO BUILD TRUST THROUGH SECURE, THOUGHTFUL AI APPLICATIONS. [4]

SALESFORCE

Trustworthy Hybrid AI is not just a smarter AI - it is a governed framework for managing enterprise knowledge and processes. Hybrid AI combines deterministic logic with governed creative GenAI to form a seamless experience.

This system starts with people choosing and organizing the right documents and client interaction steps. These are added to the company's Knowledge Graph. Trusted Collections are built using approved documents. Sanctioned documents guide GenAI to provide accurate answers with clarity. If the risk is low, GenAI responses can be shared directly with users. If the risk is higher, KMs see them first, for approval. When launched, Hybrid AI Agents offer 24/7 support through web, chat, mobile, or other tools.

The following diagram shows how the system works. It starts with gathering key knowledge and building Trusted Collections. Important facts are added to the Knowledge Graph. Then, Hybrid AI Agents are launched. Users begin to ask questions and complete tasks. The system runs as a feedback loop. Users provide input through their ques-



tions and answers. This helps find missing information or process gaps. Over time, KMs and Subject Matter Experts (SME) update and improve the system.

This isn't a one-way interaction. It's a continuous improvement system, with humans governing and improving performance at every stage. Each cycle tightens accuracy, broadens coverage, and deepens trust. Over time, the AI Agent becomes more intelligent, not by guessing, but by learning what's true from experts and user engagement.

The structured steps of the Hybrid AI Agent are:

1. Accurate documentation is curated.
2. Trusted Collections are created and vectorized.
3. Key content is loaded into the Knowledge Graph.
4. User queries enter via any omni-channel interface.
5. The Graph is checked first for deterministic responses.
6. If absent, a draft GenAI response is created from Trusted Collections.
7. User feedback is captured.
8. Human experts review and improve.
9. Graph is updated with newly sanctioned responses.
10. Over time, more queries answered deterministically.
11. Human governance remains in place.

GenAI's Sober Second Mind®

85%

OF CUSTOMER SERVICE LEADERS WILL EXPLORE OR PILOT CUSTOMER-FACING CONVERSATIONAL GENAI IN 2025. [5]

GARTNER RESEARCH

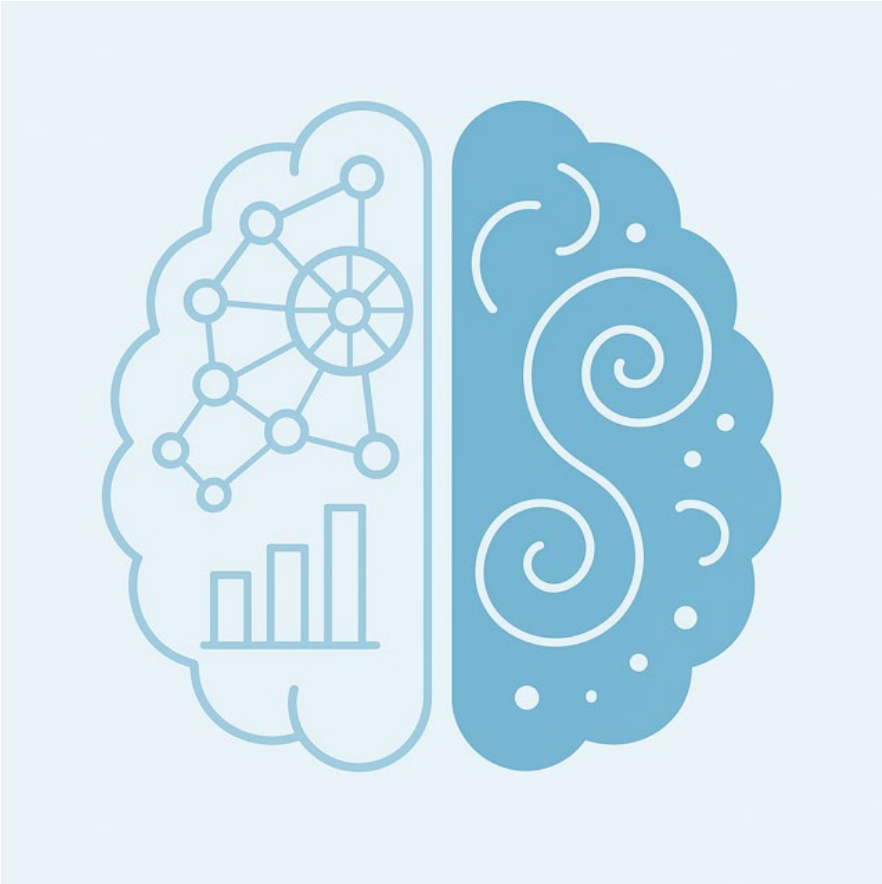
Generative AI is fast and expressive. But it should never speak for your brand unsupervised. That's why kama.ai introduced GenAI's Sober Second Mind® - to solve this problem. Think of your AI as having two sides to its brain.

The creative right brain handles drafting ideas, content, and low-risk suggestions - this right brain of the AI is the LLM. The logical left brain ensures only verified, brand-safe content reaches users. This sober part of GenAI's Sober Second Mind® is the logic engine or the Knowledge Graph AI component. It does not hallucinate and provides only human expert sanctioned answers and fully engineered process interactions supported by RPA.

GenAI's Sober Second Mind® ensures no unsanctioned GenAI output touches a customer-facing channel. This can be critically important for high-risk scenarios, emotionally and culturally sensitive circumstances, or customer-facing situations. In these cases the answers and processes the AI Agent provides are 100% sanctioned, sound, and brand appropriate.

But, this approach has the advantage of flexibility. Not all AI situations need such a tight risk tolerance where only 100% accurate answers are acceptable. For internal teams, there may be room for allowing the creativity of the right-hand brain. Your own staff will be able to use professional judgment, and are savvy to your subject matter. Here, a full Hybrid AI Agent solution works well.

This is where GenAI's Sober Second Mind® allows the LLM side to provide responses based on a trusted RAG solution. When configured to the risk tolerant mode, it can offer AI generated answers directly to the client – in



cases where a sanctioned answer is not already known by the left brain. In addition, for the KM – these generated answers and any user feedback are flagged and can be added to the 100% accurate knowledge base.

GenAI's Sober Second Mind® separates casual, helpful use cases from mission-critical AI applications. This gives you both agility and safety in a single system. Most importantly, you control the level of risk based on each audience or scenario - all while using the same enterprise knowledge that stays securely within your organization.

" 'the insidious thing' about AI hallucinations "isn't occasional errors - it's not knowing which part of the AI's answer is off" "[10]

Sridhar Ramaswamy, CEO Snowflake

Trusted Collections

As suggested earlier, Trusted Collections are a key component of the Hybrid AI Agent solution. Everything starts with the correct source material for your RAG responses. kama.ai's Trusted Collections are human-curated repositories of verified information for various domains of knowledge within the enterprise. Knowledge Administrators review and assign documents to these domains and they are then published (vectorized) in your own vector database. Careful curation of these collections avoids accidental exposure of sensitive or irrelevant data to the Trusted Collection. Responsible curation means you can't just move an entire folder into a Trusted Collection, without thoughtfully selecting each document.

At this stage the Knowledge Administrator (KA) is creating the knowledge repository for the Hybrid AI. This knowledge repository or Trusted Collection forms the basis of the generative responses that will be provided from your company's perspective. These Collections are used in two applications. First, is within the Hybrid AI. Here, RAG Draft Assist helps KMs or SMEs generate response drafts for review. These AI generated answers when approved help improve the overall system as they are inserted into the enterprise Knowledge Graph portion of the Hybrid AI Agent. Second, the generative side of the Hybrid AI powers live responses to end-user questions. This is only when the company deems it acceptable to provide end-users with direct generative responses based on the Trusted Collections.

Naturally, building Trusted Collections is important to making sure the information provided is grounded by the brand-sanctioned documents and data. In this case the Knowledge Administrator can work with SMEs to ensure

44%

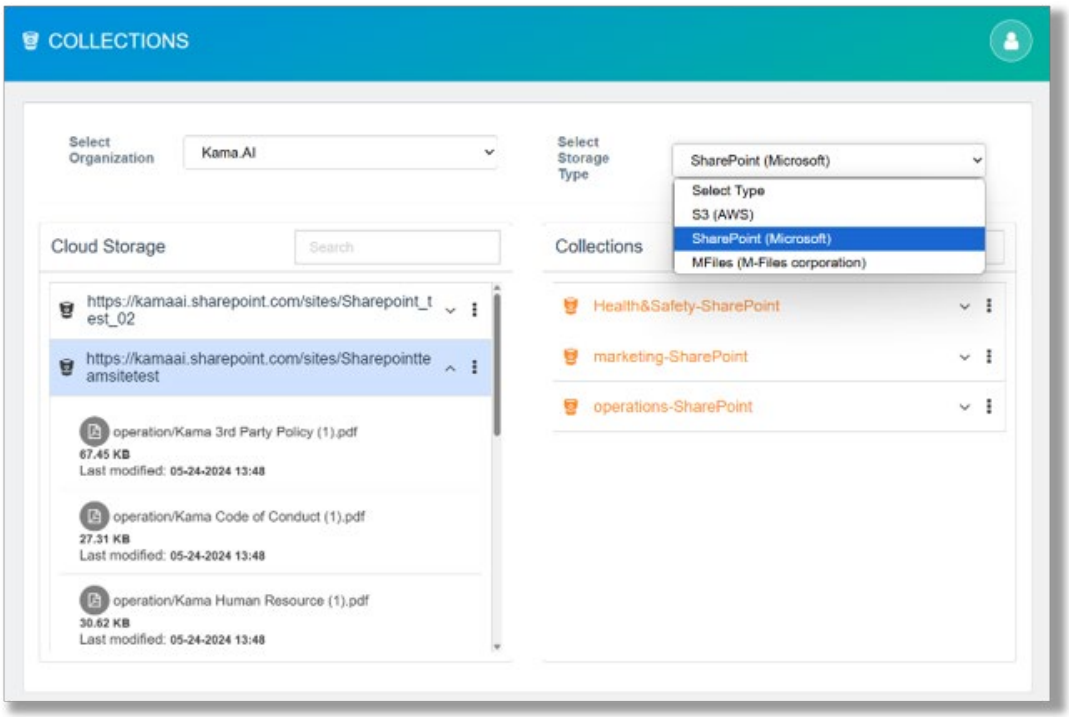
OF DEPLOYMENTS OF RESPONSIBLE AI TECHNOLOGIES CITED THE NEED TO 'HOLISTICALLY MANAGE RISK OF AI-BASED TECHNOLOGIES' AS THE UNDERLYING REASON FOR ADOPTION. NEXT WAS THE NEED TO MEET REGULATORY AND COMPLIANCE REQUIREMENTS AT 39%. [6]

PRICE WATERHOUSE COOPERS

the right information is put into the various Trusted Collections. These are based on the information domains that the enterprise wants to support.

With kama.ai's solution, Trusted Collections can be created by navigating various enterprise document repositories such as S3 (AWS), Microsoft SharePoint, or M-Files. The screenshot below shows how documents can simply be dragged from the enterprise repositories on the left, into the correct Trusted Collections on the right.

Once the Collection is finalized, it is vectorized, published and made searchable by multiple application areas within the kama.ai Designed Experience Intelligence® platform, described earlier.



Higher ROI = Responsible AI

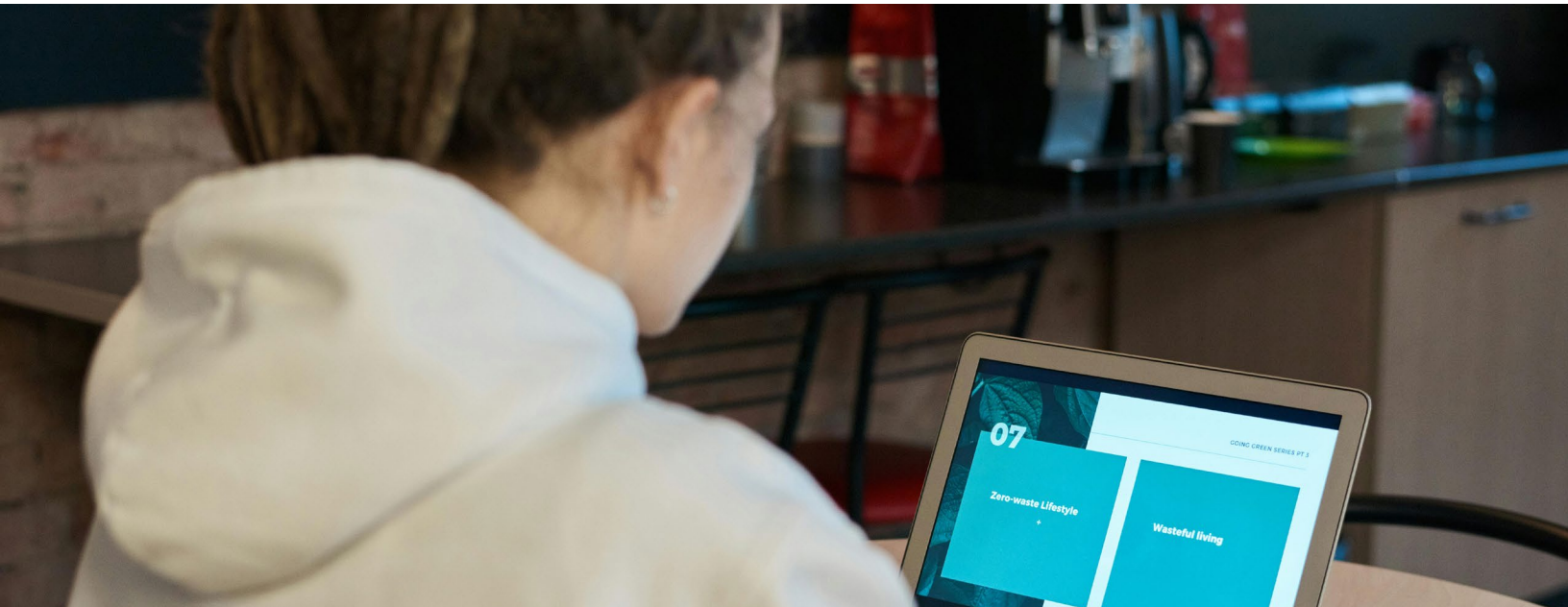
47%

HUMAN REVIEW PROCESSES REDUCE AI BIAS INCIDENTS BY 47% COMPARED TO AUTOMATION-ONLY SYSTEMS. [7]

DELOITTE

From a Responsible AI perspective Building Trusted Collections within particular domains improves the context of generative responses and reduces bias, poisoned data and hallucinations. In fact, this provides a compelling source of governance and compliance within the organization. It ensures context, consistency, and control. Beyond that, it reduces hallucination risks by anchoring GenAI in sanctioned truths – from trusted and authorized documents and data sources.

With respect to Responsible AI, the enterprise can deploy different Collections for different Hybrid AI Agents. So, within the organization each agent can be configured to deliver generative responses directly to the end user, or for more security, deliver it only to the Knowledge Administrator. This ensures semi-automated updates continue to be made to the deterministic knowledge graph or left brain of the Hybrid AI Agent. For example, you may have a customer-facing online agent that only delivers deterministic knowledge, while an employee-facing agent is empowered to deliver both knowledge graph information and generative responses based on Trusted Collections. In both cases, end users can rate responses



and provide specific feedback on information gaps.

Using Draft Assist, KMs use GenAI internally within kama.ai empowering human oversight. This ensures accuracy and efficiency. It speeds up content generation without compromising safety or truth. Teams often review Hybrid AI Agent performance responses in scheduled cycles (e.g., every Friday), using structured feedback tables and internal generative responses created behind the scenes with Trusted Collections. Full conversation histories are also available to provide context to help refine collections, documentation, and policies.

By design, this system offers:

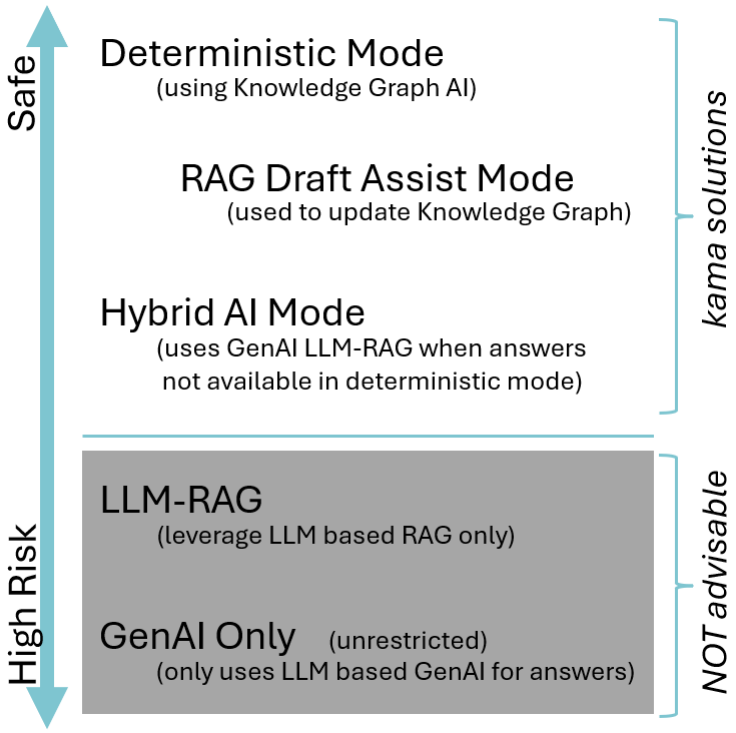
- full auditability,
- protects sensitive data, and
- ensures performance improves continuously.

It makes knowledge management and dissemination faster, safer, and more scalable. All this while protecting your brand's voice and reputation with both employees and customers.

Using Trusted Collections and Hybrid AI Agents in this way, the entire knowledge system of the enterprise is continually improved. That optimizes your return on investment through timely and ever-improving answers. An even larger benefit to the enterprise is the reduction in the effects of making the wrong decision due to a lack of information or process understanding. Plus, SMEs and management spend less time addressing repetitive questions. Instead they spend time on higher value work and strategic priorities.

Risk Profiles

Every organization has a different risk appetite. Some need 100% verified answers. Others allow more flexibility. Sometimes, it's the situation or project that defines the acceptable level of risk. With kama.ai's Hybrid AI architecture, you set how safe - or how creative - each interaction should be. The system supports structured modes aligned to defined risk profiles.



At the top of the safety scale is Deterministic Mode. Here, only pre-approved, fully validated answers from the Knowledge Graph are disseminated to the user community. Think of this as only the answers from the left deterministic side of the brain are provided. This mode gives the organization 100% accurate and pre-approved answers to questions that were already curated in the knowledge graph AI. This

is ideal for regulatory environments, legal support, public healthcare, and brand-external content. It guarantees factual precision, accurately reflects the enterprise's values, and provides audit-ready accountability.

The next level is RAG Draft Assist Mode and/or behind the scenes GenAI responses to live user chats. Both are powered by Trusted Collections. Here, the power of GenAI is used with carefully selected enterprise knowledge sources, but drafts are only visible to KMs or SMEs for review. For vectorization of Collections, kama.ai's own proprietary process is used. For the generation of responses, the enterprise can choose models from OpenAI, Google, or Anthropic. All are used safely without content ingestion into those models. In other words, the models don't get trained with your specific information or data. With some integration support, enterprises can also use their own custom model. Once a response is generated, human experts review, edit, and perfect these drafts. When approved, answers are added to the enterprise Knowledge Graph for 100% risk-free responses in future chats. The deterministic knowledge database that grows over time increases efficiency and reduces risk. This risk profile supports faster content scaling with user input and feedback. All this while maintaining full human oversight and audit control.

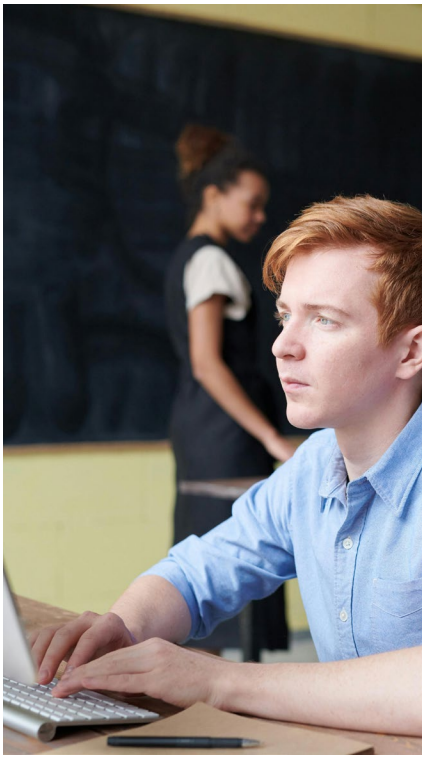
In Hybrid AI Mode, both deterministic and GenAI engines operate together through to the end user. If the deterministic system cannot answer a question, then the Hybrid AI

33%

BY 2028, 33% OF ENTERPRISE SOFTWARE APPLICATIONS WILL INCLUDE AGENTIC AI, UP FROM LESS THAN 1% IN 2024. [11]

GARTNER

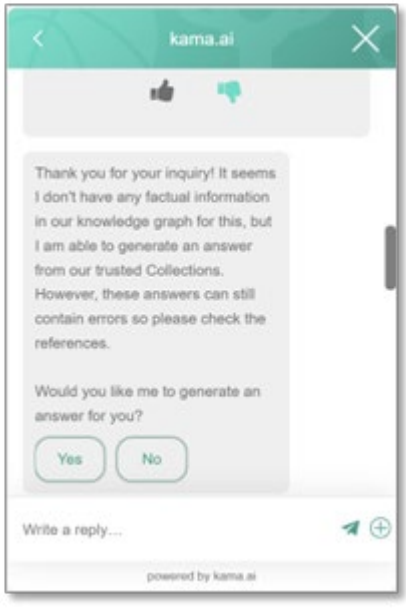
AS AGENTIC AI SYSTEMS MAKE MORE AUTONOMOUS DECISIONS AND TAKE ACTION, THEY MUST RELY ON ACCURATE INFORMATION. THIS HIGHLIGHTS THE NEED FOR HYBRID AI TO ENSURE 100% TRUSTED DATA—AVOIDING ERRORS THAT COULD HARM THE ORGANIZATION.

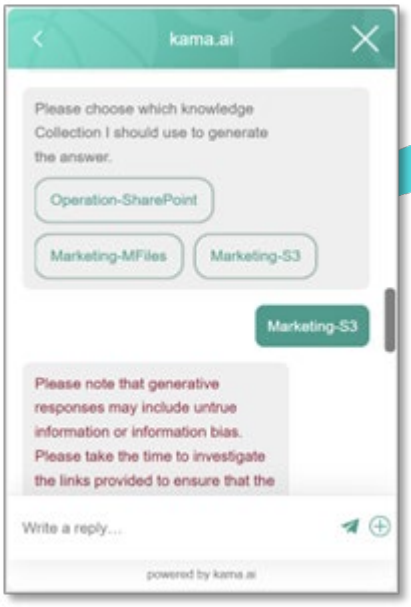


Agent architecture offers users an option to provide a response. If the user chooses a generative response, a warning is given. It states that information may not be completely accurate. This end user exchange is shown clearly in the chatbot interface images below. The interaction is transparent. Warnings are shown. Collections are named.

Beneath the threshold model lie two risky, non-recommended options: simple LLM+RAG and GenAI (LLM) Only (unrestricted). These operate without optimal grounding, strict governance, or transparency. Hallucinations, bias, and brand erosion can be expected when using these modes. This is why they are flagged as “not advisable” but are facilitated within kama.ai’s Responsible AI platform.

The structured Hybrid AI control system provided by kama.ai lets enterprises blend safety and flexibility. It enables deterministic systems for critical answers and leverages governed GenAI, grounded by Trusted Collections, for the most responsible responses — eliminating or mitigating brand risk and enterprise reputation.





Where Hybrid AI Works

Responsible and Trustworthy AI is not just for customer-facing chatbots. When designed with structured governance, clear knowledge boundaries, and human oversight, it delivers value across the entire enterprise.

In **customer service**, hybrid virtual agents resolve tickets faster and reduce escalation rates. Users get verified responses on the first try—grounded in approved policies and documents.

Sales teams use it to access up-to-date product specs, pricing, and campaign messaging. They can query the system about the differences in update features, pricing, benefits to clients, and so on. This ensures consistency and eliminates reliance on scattered internal sources or outdated assets.

Marketing benefits from enforcing a brand tone and on-messaging compliance. Most companies create a brand

15%

BY 2028, AT LEAST 15% OF DAY-TO-DAY WORK DECISIONS WILL BE MADE AUTONOMOUSLY THROUGH AGENTIC AI, UP FROM ZERO PERCENT IN 2024.[11]

GARTNER

guideline, then don't actually use it on a regular basis – to guide marketing content creation. Using a Hybrid AI solution ensures content is aligned to guidelines, and GenAI outputs can always be reviewed before material is published.

Legal and compliance teams often rely on the Knowledge Graph for policy clarity, regulatory support, contract consistency, and risk-mitigated interpretations. As an example, a contract repository can readily provide wording, critical dates, renewal notices, and so on. All these being traceable to specific internal-only documents.

HR uses the system to provide immediate answers to questions around benefits, onboarding, and DEI practices. Allowing employees access the right Trusted Repositories, helps offload the HR team from mundane question answering to their strategic work for the organization.

IT Helpdesks implement it for Tier-1 support. Agents deliver standard tech resolutions without manual triage - cutting ticket volume and response time.

Team / Function	Primary Benefit
Customer Support	Fewer tickets, faster resolutions
Sales	Smarter, compliant conversations
Marketing	Brand-safe messaging
Legal	Risk reduction, governance assurance
HR	Consistent internal answers
IT Helpdesk	Tier-1 automation
Finance	Policy clarity, consistency across team

Finance and procurement use it to enforce policy clarity and streamline approvals. Processes stay consistent, no matter who is asking.

This isn't simply automation. It's a centralized knowledge infrastructure that drives performance and safeguards trust across departments.

The chart above summarizes how each core function benefits from implementing a trusted Hybrid AI solution.

Human-in-the-Loop

3X

MORE EMPLOYEES ARE USING GEN AI FOR A THIRD OR MORE OF THEIR WORK THAN THEIR LEADERS IMAGINE: MORE THAN 70% OF ALL EMPLOYEES BELIEVE THAT WITHIN 2 YEARS GEN AI WILL CHANGE 30% OR MORE OF THEIR WORK. [8]

MCKINSEY & COMPANY

AS EMPLOYEES ARE ALREADY USING AI TECHNOLOGIES - COMPANIES SHOULD SEEK WAYS TO PROVIDE AI TECHNOLOGIES THAT HELP THEM, AND PROVIDE INFORMATION THAT IS SAFE, ON-BRAND, AND ORGANIZATION CENTRIC.

Human-in-the-Loop in probabilistic AI refers to the human engineers and analysts that tune the model and add or improve training data. This done to improve model performance which indirectly affects the responses delivered. With kama.ai's Hybrid AI platform multiple humans across the enterprise ecosystem participate by directly informing the enterprise knowledge infrastructure. This continually and directly improves AI Agent output.

Knowledge Administrators, SMEs, and KMs are the control layer. They intentionally build and maintain Trusted Collections for GenAI assistance. They also update the deterministic Knowledge Graph AI information for 100% accuracy. Nothing is added to the Knowledge Graph unless a human approves it. This means every sanctioned response carries institutional accountability. Human oversight and responsibility are essential for a well-governed enterprise AI strategy.

However, human involvement doesn't end with the internal teams. With the kama.ai platform, end users are also involved in the continuous knowledge improvement loop. Conversation logs include user inquires, direct user feedback, and automated GenAI suggestions where knowledge gaps are identified. This comprehensive approach using 'humans-in-the-loop' and generative responses based on Trusted Collections - creates a continuous feedback loop to drive improved accuracy and clarity. In effect, it continually strengthens the knowledge base for optimal enterprise efficiency, while reducing risk.

With kama's Hybrid AI Agent platform, the AI does not drift, hallucinate, or damage your brand's reputation. Human expertise governs every phase of knowledge flow:

from content development and curation to delivery and continuous improvement.

McKinsey's 2024 SuperAgency Research showed that organizations implementing AI with structured human oversight achieved 30% more operational value than those using fully autonomous systems.[8] This is not surprising. Enterprises gain more when people are engaged and in control of the intelligence layer.

Human oversight also plays a crucial role in risk mitigation. It ensures brand tone, compliance, accessibility, and alignment with organizational and user values. Most importantly, it maintains trust - not only in the system but across the organization. The AI becomes a trained extension of your people - not a rogue improviser.

In short, keeping Humans-in-the-Loop transforms AI from a content generator into a responsible brand-aligned knowledge partner. It's what makes Hybrid AI Agents trustworthy and responsible enough to represent your organization and drive an ROI for your stakeholders.

Oversight NOT Slowdown

40%

A SURVEY OF 79,000 COMPANIES ACROSS FIVE COUNTRIES REVEALED THAT 40% OF BUSINESSES CITED INACCURACIES IN AI-GENERATED INFORMATION AS A KEY CONCERN, UP FROM 14% THE PREVIOUS YEAR.. [12]

THE TIMES

Although humans are in the loop with this Hybrid AI model - it does not mean a slowdown in deployment. The human oversight ensures your company reduces its risk exposure. The Designed Experiential Intelligence® approach of kama.ai means human expert data is added directly into GenAI's Sober Second Mind®. This accelerates AI development, validation and deployment time. Governance is built into the process because information is governed in advance by SMEs and KMs. Validation is performed in a short quality approval process. As a result AI Agents can be delivered in a matter of weeks without programming or data science experts.

Continuous improvement with human oversight means the system becomes more powerful with a greater number of brand-safe answers. Each iteration reduces the need for further human intervention. Over time, this brand-safe solution lets you improve staff productivity while delivering improved response times to employees and customers. Users get the actionable information they need while employees focus on higher value add tasks and strategic projects. A clear win.

When Accuracy Matters

In some industries, accuracy isn't just preferred, it's mission critical. A single incorrect answer can carry legal, ethical, and even life-altering consequences. Here, trustworthy AI must be deterministic, governed, and fully auditable.

In banking and financial services, regulatory frameworks demand precision. A chatbot that incorrectly explains

loan terms or tax implications could result in lawsuits or fines. Fully deterministic AI ensures only pre-approved, compliant answers are shared.

In healthcare, misinformation can risk patient safety. Diagnoses, treatment plans, or benefit eligibility details need to be 100% accurate, backed by licensed medical or administrative knowledge. Hybrid AI lets you set the system to full deterministic mode but generative responses can still be delivered behind the scenes for Knowledge Manager review and consideration. This lets you deliver only reviewed content. No hallucinations. No improvisations. But, it still leverages GenAI for internal review purposes based on your Trusted Collections.

In government services, transparency and accountability are non-negotiable. Constituents expect consistent, truthful responses — especially when navigating applications, legal rights, or regulatory policy. A probabilistic AI system (like Generative AI) is unacceptable here. Trustworthy systems like GenAI's Sober Second Mind® ensures deterministic outputs, traceable sources, and clear fallback logic.

Across these sectors, the standard isn't convenience - it's credibility. When the margin for error is zero, a Hybrid AI RAG model that leverages both deterministic and generative models, is an important balance.

Scale intelligence without compromising trust.

Found this eBook Useful?

Pass it to a Colleague

Help your team stay ahead with safer, smarter AI. Forward this PDF.

Canadore College (Real-World Case Study)

In education, institutional integrity depends on providing reliable, bias-free information. Canadore College is an excellent example of a virtual agent powered chatbot set up to answer questions about programs, admissions, financial aid, and academic policy.

In 2023, the college wanted to deploy an AI virtual assistant. Earlier, they had a negative experience with another deployment that did not perform well and was difficult to maintain. Also, they were hearing mixed negative information about LLM technologies with challenges like biased responses, offensive comments, and hallucinations.

Shawn Chorney, Canadore College's Vice-President Strategic Infrastructure, Indigenous, and Learner Services, called in kama.ai for a workshop on safe and responsible AI. The Canadore team felt reassured by the knowledge based AI architecture. They especially liked the fact that only approved answers would be delivered by the virtual agent. After considering several options, they chose a dual-purpose public facing AI Agent. This agent would support prospective student questions while also providing existing students with information about college life, directions, programs, timing, and various other student related inquiries.

In this deployment, Hybrid AI Draft Assist was used extensively. However, every answer was grounded in the college's approved documentation and web content. Ultimately, every answer was human-reviewed and approved by the institution. In this case GenAI was a deployment enabler, but generative responses were not delivered directly to end users. This done to ensure student safety and to protect the college's brand reputation.

The deployment project was launched in the spring of 2024 with a ‘go-live’ target after Labor Day when students would return for the Fall semester. Canadore named their AI Agent “Panther” after their school mascot. It was successfully launched directly after Labor Day with over 300 frequently asked questions (FAQs) prepared. The performance and student experience has been resoundingly positive. Both the kama.ai and Canadore team have continuously improved Panther’s ability to respond with accurate, brand, and student safe answers.

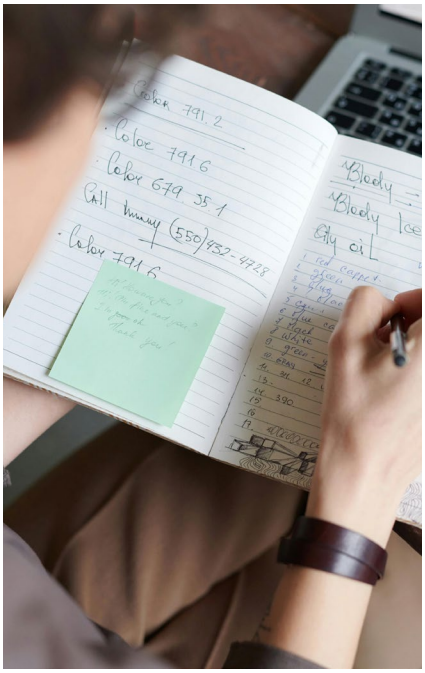
Result

Students get fast, accurate, and equitable support - day or night, weekdays, or weekends. Of equal importance, staff are freed from responding to emails and phone calls with repetitive inquiries. Rather than contending with the mundane, repetitive tasks, they now spend time on more higher value added, progressive work that more fully helps students students and the college administration.

“The intelligence behind our ‘Panther’ AI Agent is kama.ai - and it shows. We’re seeing improved response times, better student engagement, and fewer emails to manage. It’s a win all around.”

Josh Willard, Director of Marketing
Canadore College

Key Take-Aways



We explored the critical nature of trust in enterprise AI solutions. In fast-moving digital environments, brand safety, legal compliance, and information integrity can't be left to chance. Today's LLMs offer speed and fluency—but not accuracy. Here, a trusted Hybrid AI Agent platform can change the game.

Hybrid AI is about brand protection at scale. Hybrid AI gives you: deterministic accuracy with the guided agility of governed GenAI RAG. This wrapped by human experience, oversight and feedback. It isn't generic AI. It's a system built for compliance, safety, and continuous improvement.

If you remember just a few things, let them be these:

- **Accuracy at scale** matters. Even 2% error equals thousands of failed annual interactions, increasing enterprise costs.
- **Curating Trusted Collections** ensure answers reflect sanctioned, compliant knowledge - not open-web noise.
- **Hybrid AI (Draft Assist)** eliminates hallucinations. It ensures answers are generated from verified sources and governed in advance of delivery.
- **Hybrid AI (End User Delivery)** may be acceptable for some with appropriate warnings and document references.
- **'Humans-in-the-loop'** feedback, and governance turns Hybrid AI into a high performance and controlled knowledge stakeholder ecosystem - not a wildcard risk.
- **A Hybrid AI Agent** can be launched with ease while oversight makes the AI smarter without going off-brand.

GenAI's Sober Second Mind® is not a patch. It's a comprehensive platform for the safe use of AI in an enterprise. It scales responsibly. It learns from your experts. It protects your brand.

Trust is the new competitive advantage.
kama.ai delivers it - by design.

Enterprise Benefits

$14k

ENTERPRISES ARE SPENDING ON AVERAGE $14,200 PER EMPLOYEE ANNUALLY TO CORRECT AI HALLUCINATED RESPONSES.[13]

FORRESTER

The trustworthy kama.ai Hybrid AI RAG solution offers measurable value across the enterprise by accelerating deployment, controlling risk, and amplifying knowledge assets. Designed as a fully governed hybrid AI platform, it enables organizations to implement intelligent systems in just 2-3 months - without complex technical lift or unpredictable costs. This low barrier to adoption makes it ideal for teams seeking fast impact with scalable structure.

Once deployed, kama.ai enables always-on engagement - supporting both employees and customers 24/7 through intelligent, brand-safe virtual agents. The system enforces containment around the use of Generative AI, integrating with industry-leading models such as OpenAI, Anthropic, and Google Gemini, while ensuring your data never trains external models. Enterprise AI stays private, secure, and aligned to policy.

The platform also supports deeper process optimization. By reducing time spent searching, drafting, and escalating, it helps teams unlock greater productivity and make better use of institutional knowledge. Its integration with Robotic Process Automation (RPA) extends AI value into workflows - enabling automation that is conversational, contextual, and human-centered.

Beyond operations, kama.ai enhances strategic insight. The feedback loop embedded in Hybrid AI helps enterprises track user intent, customer sentiment, and internal information gaps. This intelligence - captured safely within the enterprise boundary - can guide product improvements and inform strategic priorities. In short, kama.ai enables your organization to listen to the “voice of society” while staying fully in control of what is said in return.

Let's Talk

E: INQUIRIES@KAMA.AI
M: +1 (416) 275 - 1780
W: KAMA.AI

Think you need a Responsible AI, Virtual Agent, or knowledge management solution?

Let's connect to check your options.

Ready to elevate your customer interactions with a powerful, human-centered AI? At kama.ai, we're revolutionizing the way businesses connect with their employees and clients through our Designed Experiential Intelligence® (kama DEI) platform. Our technology goes beyond typical AI - it's emotionally intelligent, personalized, deterministic (does not hallucinate) and is continually evolving with each interaction. Whether it's automating routine processes, addressing a critical employee question, improving customer engagement, or delivering 24/7 support, kama DEI brings the human touch to your AI and Virtual Agent automation.

How can better knowledge management transform your business? Let's have a pressure free conversation. Kama.ai helps organizations like yours drive sales, improve customer loyalty, and improve operational efficiency with the AI that's as human as it gets.

Think kama.ai for trust, empathy, and accuracy.

CHECKLIST

Use this checklist to evaluate AI solutions before deploying. Ensure they align with your enterprise's standards for accuracy, safety, and control. Share it with IT, compliance, or executive teams to guide responsible adoption of Hybrid AI systems.

TRUST & ACCURACY

- ○ Grounded in deterministic enterprise knowledge
- ○ No hallucinated | fabricated output
- ○ Warnings shown for GenAI responses
- ○ Responses - logged and traceable

GOVERNANCE & CONTROL

- ○ Can adjust tolerated risk level
- ○ Human-in-the-loop review supported
- ○ Safety levels set by audience type
- ○ Options for GenAI model desired

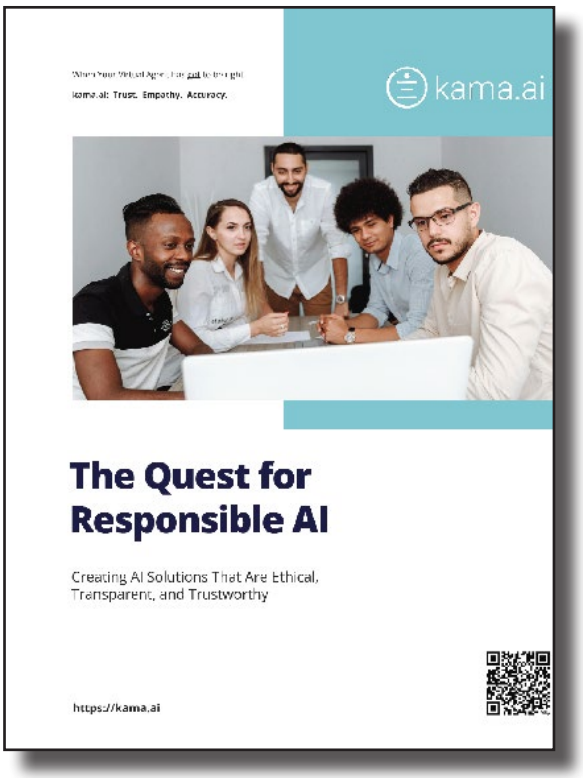
INFO & INTEGRATION

- ○ Built from company-sourced Trusted Collections
- ○ Data stays in your data repositories
- ○ Integrates with RPA & support tools
- ○ Matches enterprise tone and branding

LAUNCH & IMPROVE

- ○ Quick deploy - under 3 months
- ○ Improves from feedback and Knowledge Manager input
- ○ Works across all major channels
- ○ Low energy, ESG-friendly design

Responsible AI



Unlock the future of Virtual Agents with Responsible AI—where trust, ethics, and human values drive smarter, safer, and more meaningful enterprise interactions.

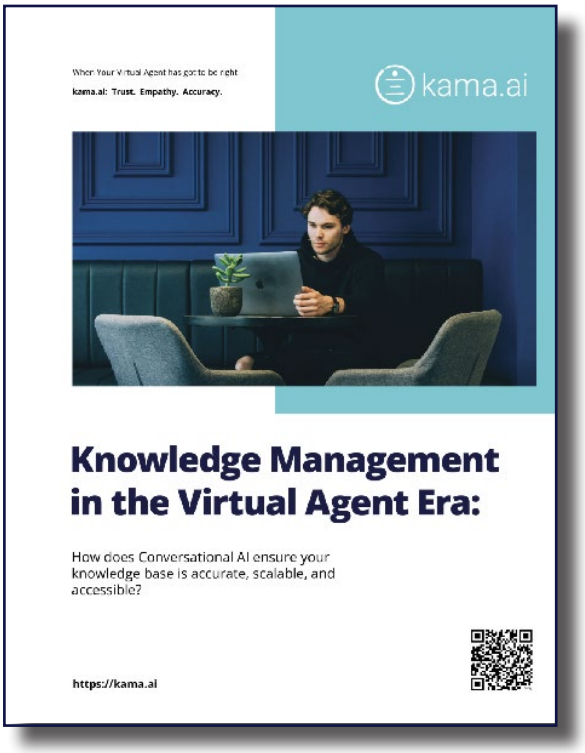
Explore how responsible AI can transform Virtual Agents into trusted, ethical knowledge partners. This ebook outlines how kama.ai's platform combines explainability, human-centered design, and governance to deliver accurate, empathetic, and values-driven information. It emphasizes the importance of transparency, privacy, and environmental sustainability in today's AI landscape.

Discover how kama.ai ensures Virtual Agents align with human values and organizational goals—building trust, protecting data, and connecting authentically with diverse audiences.

Find it here:
https://tinyurl.com/yc5vxn64

FREE to download
NO STRINGS ATTACHED !

Knowledge - Virtual Agents



A deep dive into how Virtual Agents are revolutionizing knowledge management, enhancing accuracy, and building trust within enterprise environments.

Discover the transformative power of AI in Knowledge Management in the Virtual Agent Era. This whitepaper shows how Virtual Agents, powered by AI technology, enhances knowledge delivery, efficiency, and user engagement. It details how knowledge graphs and human-in-the-loop governance ensure accuracy and align with organizational values.

Uncover the future of responsible AI-driven knowledge management. See how kama.ai's platform enables enterprises to deliver personalized, context-aware information that builds trust and connects with diverse audiences.

Find it here:
https://tinyurl.com/3b87htu3

FREE to download
NO STRINGS ATTACHED !

Endnotes

1. **Oxford University**, "Large Legal Fictions: Profiling Legal Hallucinations in Large Language Models", Journal of Legal Analysis, Vol 16. Issue 1. Matthew Dahl, Varun Magesh, Mirac Suzgun, Daniel E. Ho, Ferris Ayar, Jun 2024, https://bit.ly/4gBcd9T

2. **Journal of Medical Internet Research** (JMIR). "Hallucination Rates and Reference Accuracy of ChatGPT and Bard for Systematic Reviews: Comparative Analysis." Mikaël Chelli, Jules Descamps, Vincent Lavoué, Christophe Trojani, Michel Azar, Marcel Deckert, Jean-Luc Raynier, Gilles Clowez, Pascal Boileau, Caroline Ruetsch-Chelli, May 2024. https://www.jmir.org/2024/1/e53164/

3. **Zendesk**, "CX Trends 2025," 2025, https://shorturl.at/62vH9

4. **Salesforce**, "State of the AI Connected Customer: 7th Edition" 2025. https://shorturl.at/mPmmG

5. **Gartner Research**, "Gartner Customer Service & Support Practice." Dec 2024. https://shorturl.at/I73VU

6. **PWC**, "PwC's 2024 US Responsible AI Survey: AI is becoming intrinsic to business strategy and operations. Here's how to speed up initiatives, manage risk and generate value." 2024. https://shorturl.at/As5y9

7. **Deloitte**, "State of Ethics and Trust in Technology - Annual report: 3rd Edition." Forward by Lara Abrash, Chair. 2024. https://shorturl.at/JtvWD

8. **McKinsey & Company**, "Superagency in the workplace: Empowering people to unlock AI's full potential." Jan 2025. https://shorturl.at/i8joX

9. **Forbes Magazine**, "100 Customer Experience Stats For 2023," Blake Morgan, Senior Contributor, Jun 2023. https://shorturl.at/OQUUc

10. **Business Insider**, "Snowflake CEO explains 'the insidious thing' about AI" Nov 3, 2024. https://tinyurl.com/38tdrdch

11. **Gartner Research**, "Top Strategic Technology Trends for 2025: Agentic AI" Tom Coshow, Arnold Gao, Lawrence Pingree, Anushree Verma, Don Scheibenreif, Haritha Khandabattu, Gary Olliffe. Oct 21, 2024. ID G00818765

12. **The Times**, "Companies' fears grow about AI assistants." Artificial Intelligence, Richard Tyler, Sep 10, 2024. https://tinyurl.com/2cwbbsxs

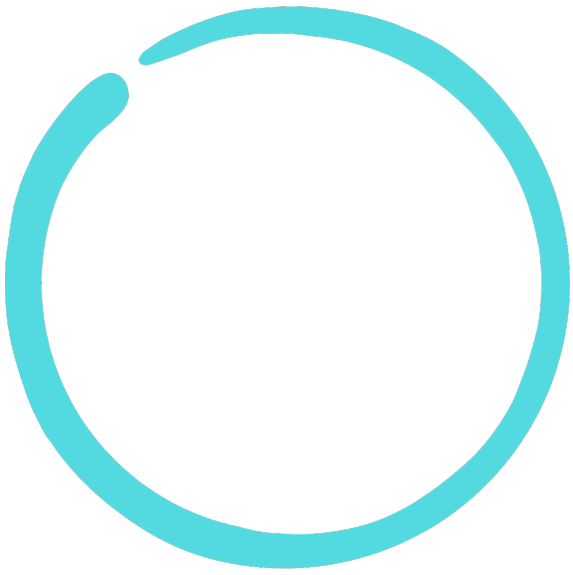
13. **Forrester** (in All About AI), "AI Hallucination Report 2025: Which AI Hallucinates the Most?" Midhat Tilawat, May 9, 2025. https://tinyurl.com/4xzt89vy

Photography Thanks To:

1. Pexel: Fauxels
2. Pexel: MidTrack

kama.ai

E: INQUIRIES@KAMA.AI
M: +1 (416) 275 - 1780
W: HTTPS://KAMA.AI



Visit us on **LinkedIn**