# Inside SARA™

## Building The First Purpose-Built, Agentic AI for Security and Monitoring

Learn How AITX's RAD and AWS's Generative AI Innovation Center Partnered to Advance Real-World AI for Security

ROBOTIC ASSISTANCE DEVICES
an AITX company

AITX
Artificial Intelligence Technology Solutions, Inc.

RAD
ROBOTIC ASSISTANCE DEVICES GROUP

## The Business of Security Monitoring Has Long Depended on Endurance.

Operators spend hour after hour watching screens, hoping to identify moments that matter. The problem is not the intention; it is the limitation.

Human concentration declines, fatigue sets in, and critical events go unnoticed. For years, this was considered an unavoidable cost of doing business.

When artificial intelligence began to enter the conversation, many believed the problem was solved. Yet the tools available were never designed for the reality of physical security. They were trained on internet images, not real environments.

They could label a scene but not understand it. They lacked precision, consistency, and accountability. The industry tried to adapt these general models, but the results were uneven and often unreliable.

## Our Approach

At AITX's RAD, we recognized that the challenge was not to apply AI, but to rebuild it for a specific purpose. True reliability could only come from a system designed around the needs of real-world monitoring: accuracy without fatigue, context without confusion, and adaptability without compromise.

> The vision for SARA began with that realization. It was not about creating another feature or incremental upgrade. It was about engineering a foundation for security that could think, interpret, and respond as effectively as a trained professional.

That goal required years of development, extensive data creation, and a team willing to design every component from the ground up.

This white paper documents that process. It outlines how SARA was conceived, built, refined, and proven in the field. It explains why purpose-built agentic AI has become the defining advantage in modern security operations, and how AITX and RAD now stand at the forefront of that transformation.

Development of SARA also benefited from collaboration with the AWS Generative AI Innovation Center, whose guidance and technical resources supported key stages of SARA's model optimization and deployment strategy.



+1.877.787.6268 | 877.78.ROBOT | www.radsecurity.com | sales@radsecurity.com

# Recognizing The Gap

Before SARA, the team at RAD spent months evaluating what the broader AI community had to offer. On paper, the technology appeared promising.

General-purpose large language and vision models could describe images, label movement, and even respond in conversational form. In practice, those tools struggled in the environments that define physical security.

The testing was revealing. Consumer and open-source AI models routinely misread situations that professionals could interpret instantly.

For example, A camera feed showing a delivery driver lingering near a dock might be flagged as “motion detected” but offered no context. A person approaching a restricted area at night might be labeled simply as “human detected” with no indication of intent, risk, or necessary response. These were not analytical systems. They were guessing engines.

That decision changed everything. The project moved from evaluation to creation.

The team began designing a purpose-built architecture that would understand context, apply reason, and deliver verified output without the instability seen in general models. The result would eventually become SARA’s visual and cognitive foundation.


## The Problem

In security, guessing is unacceptable. Every incorrect alert costs time, money, and trust. Every missed event can lead to loss or liability.

## The Solution

The RAD team realized that reliable autonomy could not come from adaptation; it required construction. To achieve the accuracy and consistency the industry demanded, the model had to be rebuilt for the specific rhythms and realities of security operations.

## Visual Foundation

# Building IRIS

Creating SARA began with building her ability to "see." Traditional AI models rely on vast public datasets, but these don't reflect the complex environments of physical security.

Warehouses, hospitals, and campuses present unique challenges in light, motion, and behavior—conditions not found in generic stock imagery. RAD's team realized that a model trained on random visual data would lack the precision needed for real-world monitoring.

To overcome this, RAD developed its own dataset creation engine: **AutoVQA**. This innovative tool generates and manages training data, allowing engineers precise control over every variable.

Synthetic captions are generated, rewritten in precise security terminology, then meticulously corrected and refined through continuous human review and performance rating. This ensures the system produces consistent, verifiable results.

---

## The AutoVQA Process:

AutoVQA's rigorous process involves several key steps to ensure unparalleled accuracy:

- **Synthetic Caption Generation:** Creating initial descriptions from video clips.
- **Security Terminology Refinement:** Translating captions into industry-specific language.
- **Expert Accuracy Correction:** Manual review and correction by security professionals.
- **Performance Rating:** Evaluating and validating against stringent quality standards.
- **Continuous Refinement:** Implementing ongoing improvement cycles for sustained precision.

# 5,000+

**Labeled captions created in the first 90 days**

The sheer volume of work was immense. Over the first 90 days, the team produced more than 5,000 meticulously labeled captions from hundreds of video clips, each manually analyzed and corrected. This was not automated convenience, but deliberate craftsmanship.

# Engineering Precision

The journey to empower SARA with accurate "vision" was grounded in a rigorous, purpose-built data generation system: AutoVQA. Recognizing the limitations of generic datasets for physical security, RAD developed AutoVQA to create an unparalleled training environment.

This meticulous process ensures that IRIS, SARA's visual foundation, understands and interprets real-world scenarios with critical precision, moving beyond mere recognition to contextual comprehension.

**Meticulously Labeled Captions in 90 Days**

This intensive manual and automated effort during the initial 90 days yielded over 5,000 high-fidelity, meticulously labeled captions.

This foundation of purpose-built data allowed IRIS to develop an acute understanding of security-critical contexts that generic models simply could not achieve.

## The Power of IRIS: Contextual Visual Intelligence

This exhaustive dataset creation process directly informed the development of **IRIS**, SARA's Intelligent Recognition and Interpretation System. IRIS isn't just about identifying objects; it's about understanding their context, relationships, and potential implications within a security environment.

It can differentiate between authorized and unauthorized access, recognize suspicious behavior patterns, and interpret subtle cues that would elude conventional AI. With IRIS, SARA gained the ability to truly "see" and interpret the world through a security lens, providing the foundation for unparalleled accuracy and actionable insights in physical security operations.

### Generate Captions

Initial textual descriptions are automatically generated from raw video footage, capturing observable events and objects.

### Rewrite in Security Terms

Raw captions are then refined by AI, translating general observations into precise, industry-specific security terminology and protocols.

### Correct for Accuracy

Human security experts meticulously review and correct each refined caption, eliminating ambiguities and ensuring factual correctness relevant to real-world threats.

### Rate Performance

Each corrected caption and its corresponding visual data are rigorously rated against stringent quality standards, evaluating the model's interpretative performance.

### Refine Through Learning

The system continuously learns from validated data, implementing iterative improvements to enhance its interpretative capabilities and reduce errors over time.

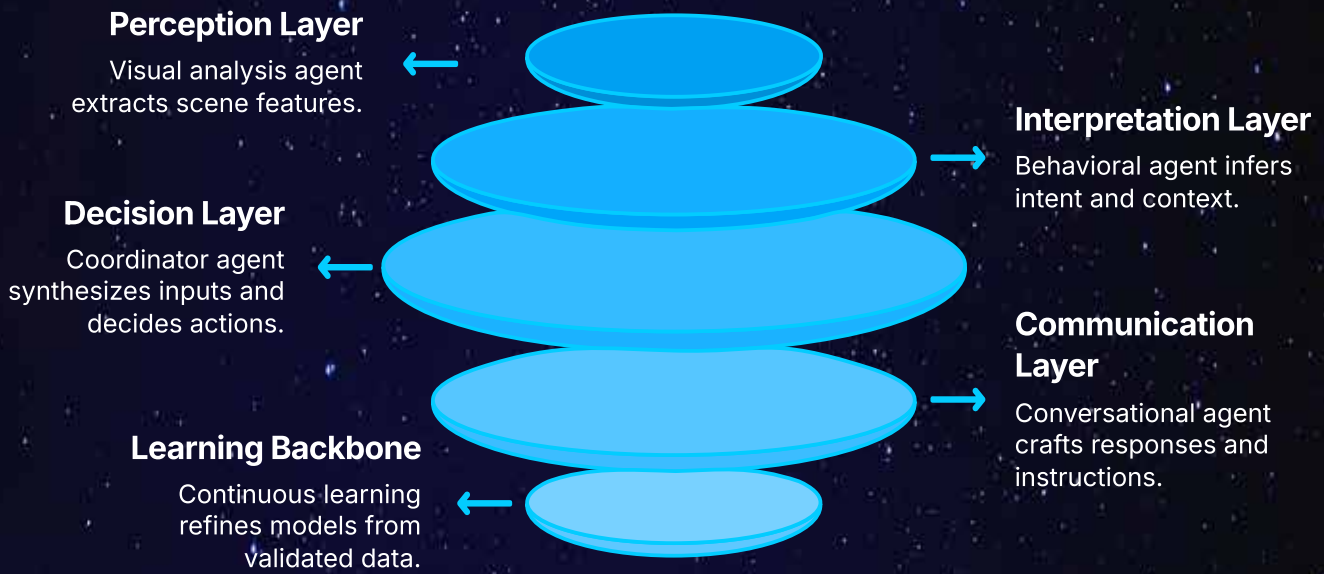


+1.877.787.6268 | 877.78.ROBOT | www.radsecurity.com | sales@radsecurity.com

# Building The Mind

Once SARA could see, she needed the ability to think. Interpreting a visual scene is only useful if the system can determine what that information means and decide what to do next. That required more than a single model. It required an ecosystem of intelligence, each part responsible for a specific function yet capable of working together in real time.

This architecture became known as **MIND**, the Multi-Agent Intelligent Network for Decisioning. MIND coordinates multiple specialized agents that operate simultaneously. One focuses on visual analysis, another manages behavioral interpretation, a third governs conversational response, and others handle data management and escalation. Each agent is trained to excel in its domain, but their true strength comes from collaboration.



# The Architecture of MIND

SARA's intelligence isn't monolithic; it's a dynamic ecosystem of specialized agents working in concert. This distributed architecture, known as MIND (Multi-Agent Intelligent Network for Decisioning), allows SARA to process complex information, interpret nuanced situations, and respond effectively in real-time. Each agent contributes a unique capability, collectively enabling a comprehensive understanding of the operational environment.

### Visual Analysis Agent

Processes raw video streams, identifying objects, movements, and environmental changes with high fidelity, forming SARA's initial perception.

### Behavioral Interpretation Agent

Analyzes detected events for patterns, anomalies, and potential intent, providing contextual understanding of actions within the security landscape.

### Conversational Response Agent

Manages all human-SARA interaction, translating complex internal insights into actionable language and processing user input for effective communication.

### Data Management Agent

Correlates information from all agents, stores pertinent data securely, refines outputs, and facilitates continuous system learning and adaptation.

By distributing responsibility among specialized agents, MIND overcomes the limitations of single-purpose AI tools. Each agent is expertly designed to excel in its domain, whether it's identifying a specific object or understanding a complex behavioral sequence. This modularity not only enhances efficiency but also allows for greater flexibility and resilience in SARA's overall operation.

Instead of relying on a singular, brittle algorithm, MIND represents a robust network where data is shared, processed, and refined through multiple layers of intelligence, ensuring a more accurate and reliable output.

The distinction between a "tool" and a "system" is critical here. While individual agents might be seen as sophisticated tools, their coordinated interaction within MIND elevates SARA to a true intelligent system. This integration allows SARA to handle unforeseen scenarios and dynamic changes in real-world environments more effectively than any standalone AI solution.

This multi-agent approach is a testament to RAD's commitment to building comprehensive, adaptable, and highly reliable AI solutions for the most demanding physical security challenges.

### From Tool to System

MIND transcends simple AI tools, creating a cohesive, intelligent system capable of complex decision-making through agent collaboration.

### Reinforced Reliability

The distributed nature of MIND ensures greater resilience and accuracy, with multiple agents validating and enriching data for robust performance.



+1.877.787.6268 | 877.78.ROBOT | www.radsecurity.com | sales@radsecurity.com

# Continuous Refinement

SARA was not designed to reach completion; she was designed to keep improving. From the start, RAD's engineering team built an internal process that allows SARA to learn, correct, and evolve without end.

Every observation, every decision, and every response becomes a new data point for refinement.

This process is known internally as the **learning loop**. Each cycle begins with a review of SARA's outputs from live and simulated environments. Engineers analyze what the system detected, how it interpreted the situation, and whether its conclusions were correct.

When discrepancies appear, they are not treated as failures but as training opportunities. The data is reviewed, corrected, uploaded, and used to retrain the model. The process repeats, tightening accuracy with every iteration.

Reinforcement learning plays a central role in this evolution. SARA receives positive reinforcement when her analysis aligns with human evaluation and negative reinforcement when it does not.

Over time, these adjustments create a system that aligns more closely with professional judgment than generic machine learning could ever achieve.

ⓘ This continuous cycle has produced measurable improvements in consistency, response time, and decision quality. It ensures that SARA's performance grows stronger with every deployment rather than degrading with age.

**Observe**

SARA monitors live environments

**Review**

Engineers analyze outputs

**Correct**

Data is refined and corrected

**Retrain**

Model learns from corrections

**Improve**

Performance increases

Unlike static AI models that require complete retraining to improve, SARA refines herself within an established framework.

Each hour that SARA operates in the field contributes to her intelligence, expanding her capabilities and sharpening her understanding of real-world environments.

> The learning loop is the quiet engine behind SARA's success. It transforms every deployment into an experiment, every correction into progress, and every interaction into knowledge.

# Engineering The Advantage

Performance is not measured only by intelligence, but by efficiency. The most advanced AI has little value if it requires excessive computing power or cannot operate reliably in the field.

From the beginning, RAD's engineers designed SARA to deliver enterprise-level capability while maintaining lightweight performance suitable for deployment across multiple platforms. Through collaboration with AWS and the Generative AI Innovation Center, RAD leveraged AWS cloud resources and optimization frameworks to refine SARA's performance profile, ensuring scalable training and lightweight deployment across multiple environments.

### The Challenge

Early training demanded significant GPU power and memory impractical for field deployment

### The Process

Model parameters streamlined, redundant layers removed, pruning and quantization applied

### The Result

Dramatic reduction in GPU utilization, operational footprint below two gigabytes

## <2GB

**SARA's operational footprint**

This breakthrough changed what was possible. It meant that SARA could function efficiently on the same hardware that powers RAD's field devices without compromising speed or accuracy. The reduced computational demand also lowered costs, improved reliability, and expanded scalability across larger networks of devices.

These optimizations were not cosmetic. They represented a critical step in transforming SARA from a research project into a commercially viable intelligence platform. It also created a framework for future expansion, where new modules or additional agents can be integrated without sacrificing responsiveness or stability.

The efficiency achieved through this engineering effort is one of SARA's defining advantages. It allows RAD to deploy advanced AI capabilities in locations and conditions that other systems cannot support, delivering consistent intelligence at scale and without compromise.



+1.877.787.6268 | 877.78.ROBOT | www.radsecurity.com | sales@radsecurity.com

# Validation In The Field

The measure of any technology is not what it can do in theory, but what it accomplishes in practice. After two years of development, SARA moved from controlled testing environments into full operational deployment across RAD's growing ecosystem of security devices. The transition marked a defining moment in the project's history.

Integration began with stationary units such as ROSA and RIO, then expanded to mobile systems including ROAMEO. Each deployment provided new opportunities for validation. In the field, SARA performed consistently, identifying potential security events, classifying behavior, and initiating appropriate responses.

Monitoring centers that adopted SARA reported measurable reductions in false alarms, faster incident verification, and improved confidence in automated alerts. AWS infrastructure has been integral in enabling RAD to train, validate, and deploy SARA efficiently across global monitoring networks.

### False Alarms

Measurable reductions reported by monitoring centers

### Response Speed

Faster incident verification times

### Confidence

Improved trust in automated alerts

The system's ability to operate autonomously while maintaining human-level oversight quickly became one of its most valued features.

SARA's multi-agent design allowed her to analyze a scene, determine the nature of an event, issue a voice-down warning, or escalate to human review without hesitation or fatigue.

What had once required multiple human operators could now be achieved through one unified intelligence. RAD's engineering and operations teams closely tracked performance metrics during this period.

The data confirmed what the design had promised: high accuracy, consistent results, and adaptability across a range of environmental conditions. These findings supported the expansion of SARA across RAD's full product line, including AVA for access control and RADCam Enterprise for small business security applications.

## Industry Recognition

The industry took notice. At the **2025 Security Industry Association New Products and Solutions Awards**, held during ISC West in Las Vegas, SARA received two of the event's top honors.

### Judges' Choice Award

The program's highest distinction across all product categories

### Best in Threat Detection

Winner for Threat Detection and Response Solutions

These accolades recognized both SARA's technical sophistication and its measurable impact in real deployments. SARA's success in the field established more than validation. It demonstrated that the architecture, training, and optimization that defined its creation were not theoretical achievements, but functional realities transforming the way security is delivered.

# The Results

SARA represents more than a technological success; it is the culmination of years of deliberate work, creative problem solving, and constant refinement.

The project demanded expertise across multiple disciplines, including artificial intelligence, computer vision, data engineering, behavioral science, and field operations. All efforts were aligned around one purpose: to create a dependable agentic AI intelligence for security.

## The Investment

Over a two-year development cycle, RAD's engineers, data specialists, and product teams built every layer of SARA's architecture internally.

The effort included designing custom dataset generation tools, constructing reinforcement learning loops, optimizing performance for deployment, and testing in live environments.

Each stage added complexity and knowledge, making the system more stable, accurate, and responsive to the unpredictable nature of real-world security monitoring.

## The Impact

The results are measurable. SARA operates efficiently across RAD's product line, maintaining precision while running on less than two gigabytes of GPU capacity.

False alarms are reduced, operator workload is minimized, and incident verification and response times are significantly faster. Monitoring centers and enterprise users report higher situational awareness and greater trust in automation.

These outcomes are not projections; they are documented through daily field use.

## Expanding Reach

In addition to its deployment across RAD's family of devices, SARA is now being developed for standalone operation. This expands SARA's reach beyond proprietary hardware and into the vast global network of ONVIF compatible camera systems.

ONVIF compliance means SARA can interface with virtually any modern IP-based security camera, regardless of manufacturer. This gives her access to tens of millions of installed systems already in use across the world, opening the potential for widespread adoption without the need for costly equipment replacement.

SARA's evolution also includes direct integration with leading video monitoring platforms used by commercial, industrial, and government security operations.

By connecting with platforms such as Immix and other enterprise-grade monitoring software, SARA becomes an intelligent automation layer capable of analyzing live feeds, identifying priority events, and initiating response actions directly within existing workflows. These integrations position SARA not just as a product, but as an enabling technology for the broader industry.

In every respect, SARA demonstrates what happens when a company commits to building technology for a singular mission rather than repurposing tools built for another domain.

The journey proved that meaningful progress in artificial intelligence requires control over the data, the model, and the purpose. For RAD and AITX, that control has become a lasting advantage, one that competitors will find difficult to replicate.



+1.877.787.6268 | 877.78.ROBOT | www.radsecurity.com | sales@radsecurity.com

# Why Purpose Built Agentic AI Wins

The security industry is moving toward a future where intelligence, not equipment, defines performance. For decades, progress was measured in hardware capacity, higher resolution, and expanded coverage. That approach has reached its limit.

The next evolution belongs to systems that understand what they see and take appropriate action without hesitation.

SARA was built for that future. Every layer of her architecture, from visual recognition to autonomous reasoning, was designed for the realities of physical security.

The result is an adaptable intelligence that continues to improve over time and can extend across devices, networks, and monitoring environments.

### Purpose-Built Focus

What distinguishes purpose-built AI from generalized models is not only precision, but intent. General models are designed to serve many industries at once, which makes them broad but shallow.

SARA is focused, trained exclusively on security data, and refined by professionals who understand how incidents unfold in real environments.

### Measurable Reliability

That focus produces consistency, accountability, and measurable reliability, qualities that generic AI cannot match.

As SARA continues to evolve, her capabilities will scale with the needs of the organizations she supports. Her intelligence will inform automation across new environments, enable faster incident resolution, and help redefine how human operators interact with AI systems.

What began as a focused response to a long-standing problem is now becoming the framework for an industry-wide transformation.

For AITX and RAD, the investment in building SARA has already proven its worth. The platform's success shows that real progress in artificial intelligence does not come from shortcuts or borrowed data. It comes from vision, persistence, and the conviction to build what the future demands. SARA stands as proof of that principle and as the foundation of what comes next.

- RAD's continued collaboration with AWS, and the Generative AI Innovation Center ensures that SARA's evolution aligns with the most advanced practices in AI training, optimization, and ethical deployment.



+1.877.787.6268 | 877.78.ROBOT | www.radsecurity.com | sales@radsecurity.com

# Key Takeaways

SARA's journey from concept to deployment illustrates a paradigm shift in physical security, emphasizing intelligence over hardware.

The following key takeaways highlight the strategic advantages and transformative potential of this purpose-build agentic AI.

## Next Evolution of Security Intelligence

SARA represents a leap beyond traditional security systems, offering advanced analytical capabilities that understand and act upon visual data, redefining performance in physical security.

## Purpose-Built AI Delivers Unmatched Precision

Unlike generic AI models, SARA's exclusive training on security data and refinement by industry professionals ensures a depth of understanding and reliability that generalized systems cannot achieve.

## Strategic Investment Creates Lasting Advantage

Two years of internal development, custom dataset generation, and continuous optimization have resulted in a robust, stable, and accurate architecture that is difficult for competitors to replicate.

## Integration Bridges to Industry Transformation

SARA's ONVIF compliance and direct integration with leading video monitoring platforms expand her reach, enabling widespread adoption and establishing her as an intelligent automation layer within existing workflows.

## Future Security: Intelligence That Never Fades

SARA's adaptable intelligence is designed for continuous improvement, scalability, and seamless evolution, ensuring that security operations remain at the forefront of technological advancement.

These principles underscore the profound impact of dedicated innovation in AI, positioning SARA as not just a product, but a foundational technology for the future of physical security.



+1.877.787.6268 | 877.78.ROBOT | www.radsecurity.com | sales@radsecurity.com

# Pioneering The Future of Security With SARA

The journey of SARA, from concept to deployment, underscores a fundamental truth in the evolving landscape of artificial intelligence: true innovation in critical sectors like security comes from purpose-built design. Rather than adapting generalized AI models, SARA was meticulously engineered for the unique demands of physical security, setting a new benchmark for precision, reliability, and actionable intelligence.

SARA's achievements are a testament to this focused approach. Her seamless integration with tens of millions of existing ONVIF-compatible cameras, combined with direct links to enterprise monitoring platforms like Immix, demonstrates a profound understanding of industry realities.

This capability allows SARA to analyze live feeds, identify critical events, and automate response workflows with an accuracy and efficiency that generic systems simply cannot match. It represents a paradigm shift from passive observation to proactive, intelligent action.

### Scalable & Adaptable Intelligence

SARA's architecture is engineered for continuous learning and evolution, ensuring her capabilities expand to meet new threats and operational demands across diverse environments.

### Enhanced Human-AI Collaboration

By automating routine tasks and highlighting critical incidents, SARA empowers human operators, freeing them to focus on strategic decision-making and rapid, informed responses.

### Redefining Security Paradigms

SARA is not merely a tool but a foundational technology, driving an industry-wide transformation towards more intelligent, proactive, and resilient security monitoring systems globally.

Looking ahead, SARA is positioned to be the framework for how physical security operations will function. Her intelligence will continue to inform automation, enable faster incident resolution, and fundamentally redefine how human operators interact with sophisticated AI systems. For AITX and RAD, SARA is more than a product; she is a strategic advantage, a commitment to building a safer, more secure future powered by truly agentic AI.

---

To learn more about SARA's capabilities and how she can transform your security operations, please visit our website or contact our team for a personalized demonstration.

Discover SARA | Contact Us



+1.877.787.6268 | 877.78.ROBOT | www.radsecurity.com | sales@radsecurity.com