LatentAI

WHITE PAPER

AI Under Fire

Designing Edge AI Systems to Survive Mission Reality

Table of Contents

INTRODUCTION

AI in the Chaos of Combat

For AI-enabled defense, speed is a given. The force that deploys first holds an advantage.

This AI sprint is playing out in real time for companies developing the next generation of models. In early 2026, the U.S. Secretary of War announced that "a primary procurement criterion for future AI model acquisition" is that new models can be deployed within 30 days of public release. Simply stated, "speed wins."

Military advantage hinges more than ever on accelerating AI deployment. But speed is only part of the story. At the tactical edge, warfighters don't care how quickly an AI model came to production. Here's what they do care about: whether their systems can adapt to the unscripted chaos of combat.

Beyond the controlled lab, AI systems fail most of the time, not because the models are wrong, but because they were never designed to survive reality. The world at the edge is more complex, constrained, and adversarial than any test environment. It's where AI stops being a demo and becomes a dependency.

> "If these systems cannot be validated against the harsh realities and complexities of conflict, there is a risk of building a force that does not make better decisions but simply makes bad decisions faster"
>
> -Benjamin Jensen and Yasir Atalan in CSIS

Battlefield systems need to survive and create decision advantage in extreme conditions, where nothing is controlled. They need to withstand nature, physics, and adversaries without the convenience of the cloud or any lifeline to command-and-control or machine learning (ML) engineers. At the edge, AI systems and AI-equipped warfighters need to stay operational at an entirely new scale.

These are exactly the problems we solve at Latent AI, and our team has developed standard practices over the years to build adaptable, field-ready AI. In this paper, we examine:

- Why adaptability is a central goal for battlefield AI
- How to design adaptive systems for dynamic mission requirements
- How we look beyond models and approach edge AI survivability as an AI systems engineering challenge

We'll share a snapshot of our methodology and offer guidance for technology and mission leaders who are exploring AI capabilities at the edge.

Four Standards for Edge AI Survivability

As we see in the ongoing conflicts in Ukraine and the Middle East, modern warfare is moving too quickly and unpredictably for "set it and forget it" AI capabilities. The side with tactical advantage is the one that can rapidly adapt tactical edge AI systems at scale and in near real time in the field.

"Acquiring sufficient data that is diverse enough to be useful is tough," a former Pentagon official who supported Project Maven shared with CSIS. "We have a labeled dataset that is good for desert clear-sky, high-sun conditions, but if you take that AI model to a snowy landscape, the performance drops."

Latent AI wrestles with this unscripted reality every day. As part of our work for the Navy's Project AMMO, we helped our customers adapt to shifting environments on land and at sea. "The bottom of the ocean in the Red Sea looks different from the bottom of the ocean off Hawaii," said Alex Campbell, Navy service lead for the Defense Innovation Unit (DIU). "Your AI has to adapt to a different ocean and to constantly changing adversary tactics and adversary capabilities."

We know that AI system adaptability is a key metric of survivability, but it's not always a rigorous engineering metric.

What traditional ML metrics don't account for

Across the AI industry, useful metrics abound for model reasoning, accuracy, and recall. For example, benchmarks like Massive Multitask Language Understanding help test and compare LLM model performance as a critical step before deployment. For models that need to track, detect, and target, metrics like Mean Average Precision (mAP) measure how accurately they find the right objects in the right places. Often, these model metrics and benchmarks meet the essential requirements for commercial or academic AI applications.

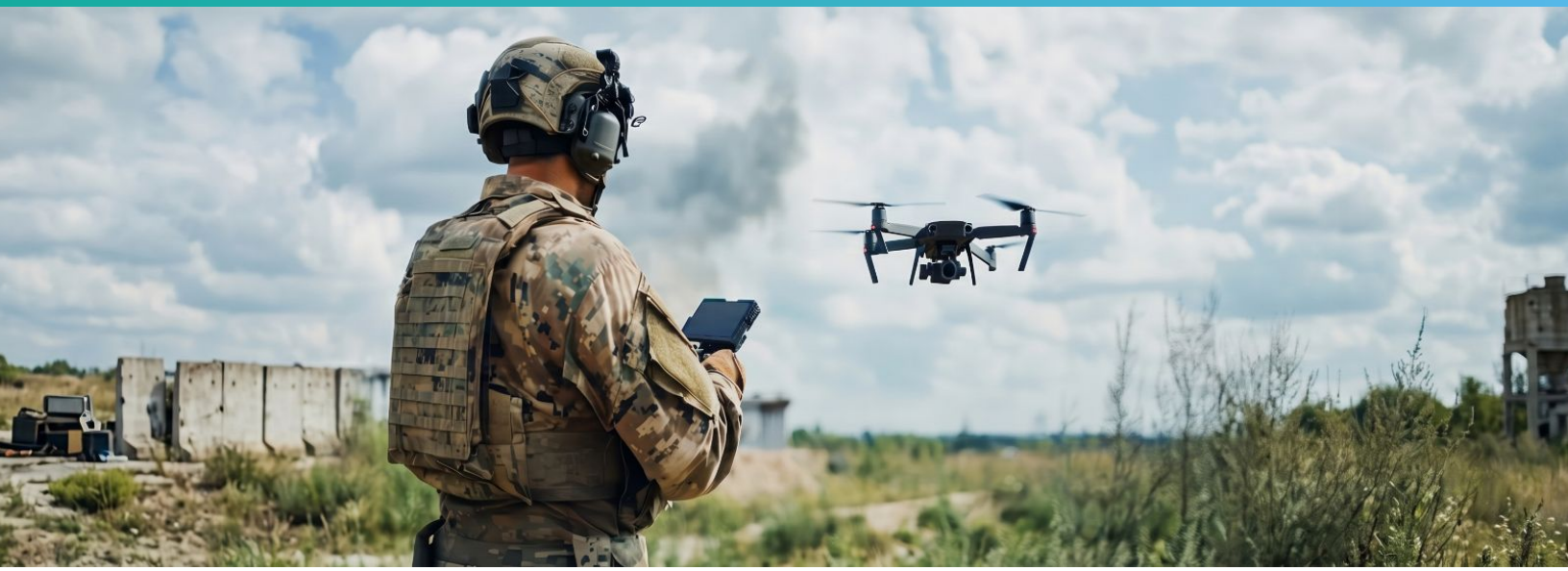
For the defense mission, and more generally for edge AI, metrics and benchmarks that test model performance absolutely matter before deployment. But even with those measurements, there is still a massive gap between model testing and battle realism.

When it comes to critical operations at the edge, here's what conventional and controlled metrics **don't** account for:

- Will users be able to adapt the models reliably in the field?
- Will the system work in extreme environments?
- Will users trust the AI in the heat of the mission?
- What happens when an adversary disrupts the environment or the system?

At first contact, assumptions collide with reality. So, what's required to ensure edge AI adaptability outside the stability of the lab?

Four Standards for Edge AI Survivability



System requirements that help AI survive the real world

Today, adaptable AI isn't theoretical. At Latent AI, we engineer for adaptability standards from day one, long before deployment. Here are four basic requirements that define edge AI survivability and separate our edge AI systems from "good on paper" to mission-ready.

REQUIREMENT #1: AGILITY AND USABILITY

WHAT WE TEST FOR: **Can users at the edge easily and safely tune the system?**

WHY IT MATTERS:

If an AI system can't be retrained or retasked in the field, it risks becoming obsolete. Today's warfighters need the ability to adapt models on the fly, without breaking them, and without an ML engineer in sight. For example, what if an adversary simply changes the paint color on a tank to mislead common assumptions? By engineering for adaptability, AI models will keep up with the threat landscape, retune themselves with ongoing data uploads, and evolve as quickly as the mission.

OUR APPROACH:

We build intuitive interfaces with options and pathways that we test and certify with customers before deployment. This way, operators can run and refine AI models in the field to adapt to threats or new data, without the need for connectivity. With the rigor of technical guardrails and robust security, our systems are built to fully empower soldiers and non-ML experts to re-tune AI in the field for a decisive advantage.

Four Standards for Edge AI Survivability

REQUIREMENT #2: EFFICIENCY

WHAT WE TEST FOR: Can the model run efficiently across different platforms and environments?

WHY IT MATTERS:

Just like all software, if AI is locked to one hardware configuration, it will fail when that configuration changes. The same is true if an AI system is exclusively "edge" or "cloud." By engineering for efficiency (vs. rigidity to one platform or to one compute environment), battlefield systems get the most out of hardware, use fewer resources, and more effectively adjust to battlefield conditions.

OUR APPROACH:

As part of designing models for efficiency, our team optimizes hardware and compute constraints, from memory to power. This requires observing and measuring AI model performance tied to cloud and edge GPUs, on all target hardware, and when network conditions evolve. Even if platforms and environments change, we build AI systems that operate continuously, securely, and predictably in the field.

REQUIREMENT #3: ROBUSTNESS

WHAT WE TEST FOR: What is the most precise algorithmic mix for target conditions?

WHY IT MATTERS:

Relying on a single, universal model for all scenarios often leads to failures during edge cases at the tactical edge. For instance, a system optimized for daytime desert operations will likely struggle in urban or nighttime environments. Furthermore, forcing a single model to handle every situation can degrade its overall accuracy while unnecessarily increasing its computational size and power consumption.

OUR APPROACH:

We develop AI algorithms that leverage situational context to deliver superior results. By factoring in variables like weather conditions, terrain, and illumination, we optimize our models for their specific environments. Similar to attention mechanisms, we orchestrate ensemble models to maximize both algorithmic performance and processing efficiency.

Four Standards for Edge AI Survivability

REQUIREMENT #4: INTEROPERABILITY

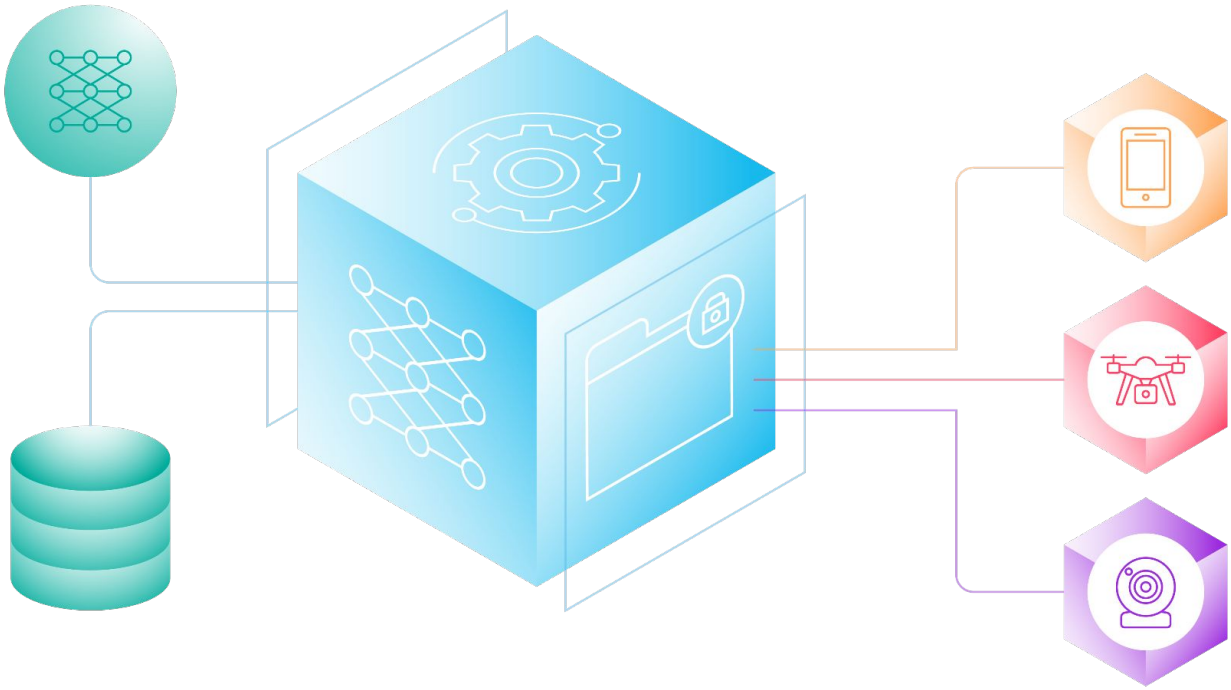
WHAT WE TEST FOR: Does the AI system offer flexibility in deployment?

WHY IT MATTERS:

When organizations are locked into limited deployment pathways, the AI supply chain becomes structurally impractical to maintain, adapt, and scale. We urgently need more flexible options to deploy AI systems across varied hardware and environments. By engineering edge AI with federated, interoperable design principles, we can break free from these rigid constraints, ensuring that mission collaboration seamlessly spans national and global coalition defense efforts.

OUR APPROACH:

By fully embracing a Modular Open System Approach (MOSA), our engineers design edge models that prioritize true interoperability, supply chain flexibility, and platform independence to eliminate single points of failure. We rigorously test across formats, hardware, interfaces, and security to resolve friction before deployment.



Model-to-Hardware Optimization

Warfighters need to adapt to all conditions: freezing, wet, offline. So do their tactical edge AI systems, which include the full stack of software and hardware.

That's why meeting AI model standards is only the first gate. Our adaptive edge AI solutions are not only optimized for the specific hardware they'll run on but also proven in real-world, physical conditions. Without a systematic approach to both, teams are left to discover failures on the physical system itself: too costly, too slow, and for organizations like the Pentagon, completely impractical.

Blending models with systems engineering, minus the guesswork

Different mission applications have different priorities and unique design tradeoffs. Targeting systems must be *accurate* above all else; surveillance systems must be *fast*. Other system requirements—power efficiency, computational availability, or latency demands—further constrain how a model can be implemented on a target device. Across the industry, these requirements can extend go-to-market timelines by up to 12 weeks. At scale, with multiple edge devices and separate optimization pipelines, the complexity becomes exponential.

The traditional approach makes this worse: ML engineers build the best, most accurate model and throw it over to the embedded hardware team, who squeeze the model into constrained hardware, cutting corners and making tradeoffs to meet requirements. Engineers aren't working from shared design principles; they're working in a vacuum.

That's why we built an active learning approach that treats model-to-hardware optimization as a design process rather than an afterthought.

> After years of deployments, we were sitting on **12TB of real-world telemetry data** from over **200,000 device hours**. We applied that rich knowledge base to build more than 1,000 pre-tested, curated model-hardware combinations, a baseline that streamlines the entire edge AI lifecycle across diverse hardware, from drones to sensors.

Instead of spending weeks trying different configurations, developers can identify optimized model-to-hardware recommendations in mere hours. Every new test and deployment feeds back into the knowledge base for continuous improvement.

"To compare performance systematically, we score combinations across criteria and target metrics," explained Sek Chai, Latent AI's CTO. "When teams start testing different blends, I think about our process like pulling ingredients for a recipe. You measure the salt, the flour, the sugar, and can see exactly how each ingredient affects the outcome. This way, testing becomes easy to replicate quickly and at scale."

PART 2: A LOOK IN THE LAB

Model-to-Hardware Optimization

From the beginning of every project, design requirements serve as a North Star. ML engineers build and tune models with clear system parameters in hand, then systematically compare and curate software-hardware combinations across platforms, rather than optimizing in isolation and hoping for success.

Why we still fly drones into the ground

A rigorous pipeline reduces the search space, but it doesn't eliminate the unexpected. Simulated testing isn't enough when the stakes are high, which is why Latent AI maintains a hardware lab where multidisciplinary teams embed AI onto physical systems, weigh performance, and deterministically identify the most optimal combinations for mission parameters.

Those tests don't always go smoothly.. Jon Brookshire, an ML engineer at Latent AI, described a recent test flight:

"Something went wrong in the control filter, and the drone suddenly thought it was falling. It accelerated itself and shot up 200 feet in the air and then crashed to the ground. We certainly considered that a 'failure,' but it was a necessary one, because it shows how many unexpected things happen when you put autonomous software into hardware."

Field validation isn't open-ended exploration, because by the time a model goes to hardware, the pipeline has already narrowed the design space. Testing is targeted confirmation, not discovery. And when something unexpected does happen, it sharpens the knowledge base and strengthens every deployment that follows, instead of sending teams back to square one.

That's how Latent AI cut edge AI development time from 12 weeks to a few hours, without skipping the hard part.

Core Principles that Govern Adaptable AI

As the Pentagon moves to “usher in an unprecedented era of American military AI dominance” and agencies respond to high expectations for AI initiatives, defense partners will need to balance development speed with the rigor of readiness testing. Because one thing worse than a slow innovation cycle is a solution that collapses when it meets reality.

That’s why the ultimate benchmark for AI at the edge is adaptability.

What AI adaptability looks like in action

Earlier, we mentioned one environment where literally nothing is constant: the bottom of the ocean. A few years ago, it took the Navy six months of painstaking human analysis, aided by computer vision AI, to clear commercial shipping zones and contested waters. The service needed an easier and faster way to track, modify, and adapt models rapidly at a massive scale.

Through Project AMMO, the Navy worked with Latent AI and other partners to build a trusted infrastructure and ML pipeline designed specifically for adaptability.

One of the most significant outcomes from Project AMMO was reducing model update time by 97%, collapsing what was a slow, intensive process into something fast, repeatable, and operationally viable.

This wasn’t just an efficiency gain. By dramatically shortening the update cycle, AMMO now allows systems to maintain a portfolio of optimized models and continuously select or deploy the one that best fits the moment: higher precision when compute is available, lower power when endurance matters, and faster inference when latency is critical. The result is an AI capability that doesn’t just perform well under fixed conditions, but keeps working as those conditions change, which is ultimately the difference between a model that impresses in testing and a system that survives in the field.

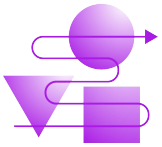
Core Principles that Govern Adaptable AI

Design principles that govern our approach

At Latent AI, we don't think about adaptability as a product feature. It's a core design philosophy that our team has operationalized through experience. Here are the principles that govern our own work, which we believe can be applied to other platforms and services across the industry:

Approach edge AI as a systems engineering challenge.

In practice, getting AI ready for the unpredictable world isn't as much of a "model problem" as it is a systems engineering problem. Every time we bring AI solutions from the lab to production in the field, we tackle edge AI survivability as a systems-level challenge, spanning the model, application, and device.



Engineer for adaptability from the start.

The goal isn't just faster, more efficient solutions, but also adaptable AI that degrades gracefully, updates securely over the air, and remains operational in challenging scenarios. In a domain where most models fail the moment things change, Latent AI invests in open, robust, and interoperable systems explicitly designed to keep working regardless of hardware, connectivity, or environmental conditions.

Create a rich foundation of knowledge to build on.

Simulated testing is never enough to assess adaptable AI at the tactical edge, but extended go-to-market timelines are, unfortunately, prohibitive at scale. To accelerate development but also refine the physical testing process, never start from scratch. At Latent AI, this philosophy (and our curated knowledge base of model-hardware blends) allows our engineers to focus on mission requirements and not spend time solving old problems.

Core Principles that Govern Adaptable AI

Edge AI that's trusted and operational under fire.

Today's warfighters don't need AI that scores high on controlled tests or that is optimized for speed alone. They need adaptive edge AI solutions intentionally designed to survive and create value amid the chaos of action.

At the edge, variables like speed, precision, and power are just tradeoffs. *Adaptability is the whole point.*

About Latent AI

Latent AI builds edge AI systems designed to survive mission reality. Where most AI solutions are optimized for controlled environments, Latent AI engineers for the edge from day one across the model, hardware, update pipeline, and deployment architecture.

Our work spans Army and Navy programs where adaptability is not a design preference but an operational requirement. Through programs like the U.S. Navy's Project AMMO, we have demonstrated what it means to build AI that not only deploys but also sustains, updates, and performs when conditions change.

We bring 200,000+ device hours of real-world telemetry and a hardware-agnostic runtime to every engagement, giving defense and enterprise customers a foundation that others are still building toward.

Contact Us

For more information, please email info@latentai.com visit **latentai.com.**