White Paper

Confidential Computing: Powering the Next Generation of Trusted AI

This paper explores Intel's strategy and solutions for improving security in the growing Generative AI (GenAI) market with Confidential Computing.

Authors

Paul O'Neill
Senior Director, Confidential Computing, Intel

Matt Hopkins
Business Strategist, Intel

Jesse Schrater
Principal Engineer, Data Center & AI Security, Intel

Table of Contents

Executive Summary

Organizations across a broad range of industries recognize Generative AI's transformative potential and have made it a top priority. And now these companies are turning to their proprietary data—their strategic moat—to extend the capabilities of AI beyond the foundational models and apply it to new use cases that maximize the value of improved data insights.

However, data integration challenges, privacy concerns, and security risks can slow down or derail companies' GenAI pursuits. These obstacles are particularly acute for companies operating in regulated industries, relying on legacy systems, or managing hybrid computing environments.

To help organizations realize GenAI's benefits without compromising their risk posture, Intel's Confidential Computing technologies offer continuing innovations like encrypted offload to accelerators and quantum-safe cryptography. Combining this with open software and a robust ecosystem helps customers:

- Protect enterprise data and models throughout the AI workflow.
- Ensure security and privacy from the edge to the cloud.
- Use legacy and emerging IT systems.
- Leverage the latest and greatest AI hardware.

The Rise & Threat of Generative AI

Despite its recent emergence, more than half of companies are estimated to have GenAI workloads in production . These companies are investing in GenAI use cases that span business functions—from drug discovery and risk management to real-time customer support and personalized marketing—hoping to drive down costs, improve insights, and transform their businesses.

Many companies, especially large enterprises, could unlock significant value from their vast data estates by enabling more effective and accurate AI with methods like fine-tuning and augmented retrieval. In the GenAI era, data will serve as a competitive moat.

However, security risks, data silos, and compliance concerns limit companies' ability to leverage these data and capitalize on their competitive advantage. As an emerging technology, GenAI introduces new security and privacy risks that companies must navigate, from long-standing concerns like data theft to novel issues like hallucinations, prompt injections, and data poisoning.

Failure to mitigate these risks could cause meaningful reputational, financial, and operational damage, as well as the opportunity cost of missed transformations. Recognizing these potential impacts, companies cite security and privacy concerns as a leading obstacle to deploying GenAI use cases.

In his blog on Creating a Foundation for End-to-End AI Security Solutions, Jesse Schrater points out that AI is transforming both the scale and predictability of data-driven decision-making. Where once we had clear control over inputs and logic, we now face vast, unstructured data sources and non-deterministic models that evolve in ways we can't fully predict or audit. Traditional security approaches—like static defenses and perimeter controls—are no longer sufficient. As AI logic operates across billions of devices and cloud platforms, security strategies must evolve to match its dynamic and opaque nature, ensuring innovation doesn't outpace the ability to govern it.

Introducing Confidential AI

Intel's Confidential AI capabilities arise from its Confidential Computing products, Intel® Trust Domain Extensions (Intel® TDX) and Intel® Software Guard Extensions (Intel® SGX). These technologies use a hardware-based trusted execution environment (TEE) to protect sensitive data and applications from unauthorized access.

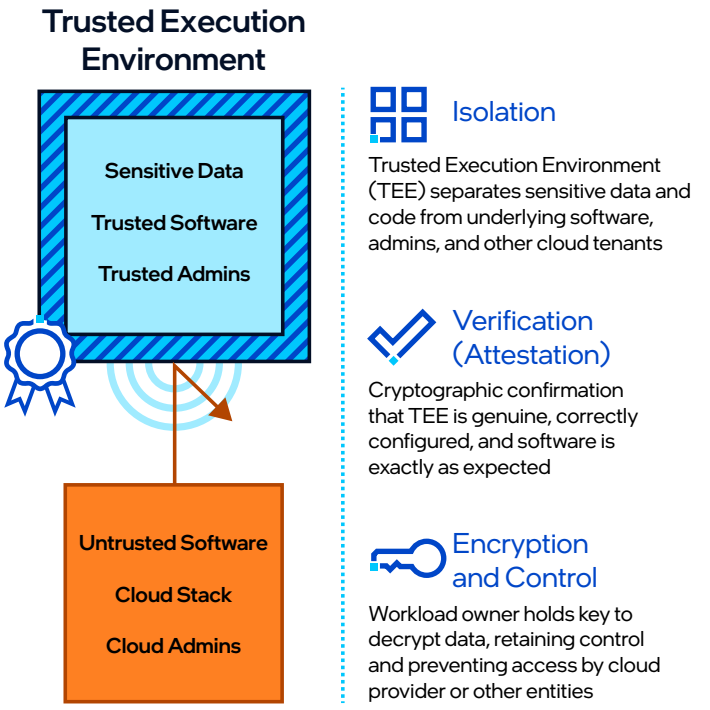


Figure 1. A Trusted Execution Environment prohibits unknown and unwanted infiltration from unverified sources

When applied to AI workloads, Intel's Confidential Computing products help companies overcome the security and privacy concerns that inhibit them from extracting value from their data and GenAI. These benefits extend from the edge to the cloud and support data stores—new and old—with the hardware required for the job.

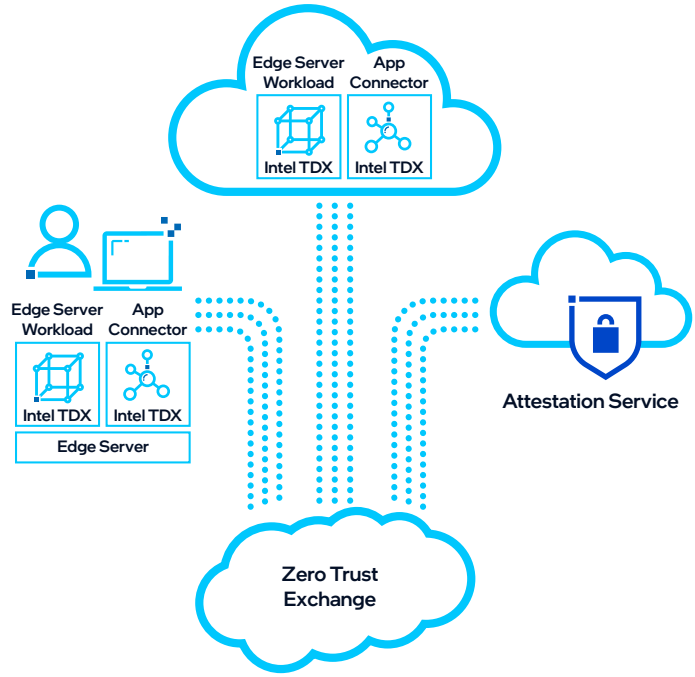


Figure 2. Intel TDX allows verification of end users and cloud sources in a zero trust exchange

Every AI workload, whether in the cloud or on a device, can benefit from a secure environment that is tailored to fit its specific needs. Intel's Confidential AI has the flexibility to support the emerging "AI everywhere" world.

1. For deployments where offload accelerators are not required (e.g., small inferencing models, RAG VectorDBs) or are impractical (e.g., TCO factors), CPUs with Intel TDX and Intel SGX support Intel® Advanced Matrix Extensions (Intel® AMX) accelerations to improve AI performance.
2. For deployments requiring an accelerator or GPU, Intel is introducing Intel® TDX Connect[1] to provide a secure channel for communicating directly with PCIe-compliant accelerators from any vendor.
3. As AI moves to the edge, Confidential Computing provides workload and platform verification (attestation) to support Zero Trust assurances for endpoints.

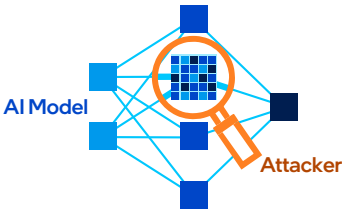
1. As a stepping stone to the first phase, Intel supports the secure use of Nvidia accelerators using bounce buffers—a software-based solution that has some performance overhead but can provide core functionality ahead of availability of the full hardware-based and performant Intel TDX Connect.

Threats to AI

Security researchers and organizations, from the Open Worldwide Application Security Project (OWASP) to the National Institute of Standards and Technology (NIST), have documented the threats that GenAI systems face. Protecting these systems typically calls for a layered security approach, with Confidential Computing supporting risk mitigation efforts.

Data Theft / Disclosure:

GenAI models rely on vast datasets, which often include sensitive or proprietary information—sometimes inadvertently (e.g., camera images with faces, license plates, etc.). During model training or inference, this data must be decrypted and loaded into system memory, making it vulnerable to attack if not properly secured. Threat actors targeting these stages can potentially extract confidential inputs, outputs, or even model parameters, leading to regulatory exposure, customer mistrust, Intellectual Property (IP) theft, and reputational harm.

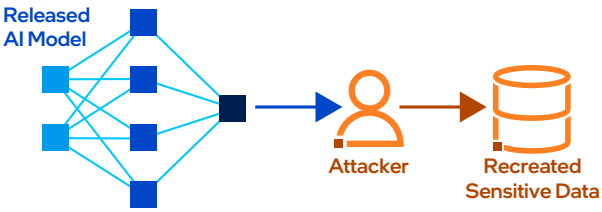


Mitigations:

Confidential Computing protects sensitive AI data by encrypting memory and isolating processing environments, helping ensure data remains secure even while actively in use. This is especially critical across the AI pipeline—during model training, fine-tuning, and inference—when raw data and intermediate representations are most exposed. By running AI workloads inside hardware-based Trusted Execution Environments (TEEs), organizations can safeguard proprietary models and private user data from malicious actors, infrastructure administrators, or compromised system components.

Model Inversion Attacks:

These attacks exploit access to a machine learning model and its predictions to reconstruct sensitive elements from the training data. In generative and predictive AI systems—particularly in healthcare, finance, or personalized services—this could mean recovering private patient records, biometric data, or other personally identifiable information. Attackers leverage the model's learned patterns to reverse-engineer inputs, making even seemingly benign API access a potential privacy risk.

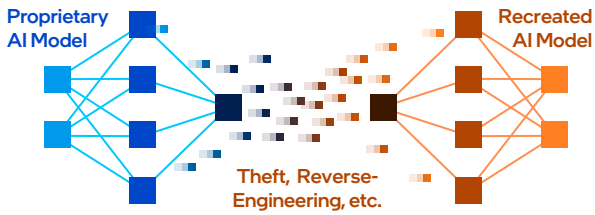


Mitigations:

Confidential Computing helps defend against model inversion attacks by enforcing a secure, attested execution environment for AI workloads. Within an Intel-based TEE, access to the model can be strictly controlled and monitored. A trusted, attested inference module can restrict the types of queries allowed, prevent excessive or abnormal interaction patterns, and apply privacy-preserving techniques (like differential privacy or query throttling) before releasing responses. This containment is designed to ensure that only verified software with a limited interface can interact with the model, dramatically reducing the risk of data leakage or inversion.

Intellectual Property Theft:

Proprietary AI models represent a major investment in data acquisition, engineering, compute, and domain expertise. These models—whether foundation models, fine-tuned LLMs, or specialized neural networks—are core intellectual property. If exposed, they can be reverse-engineered, cloned, or illicitly redistributed, enabling competitors to replicate capabilities without incurring the original cost, eroding competitive advantage, and undermining trust in AI deployment.

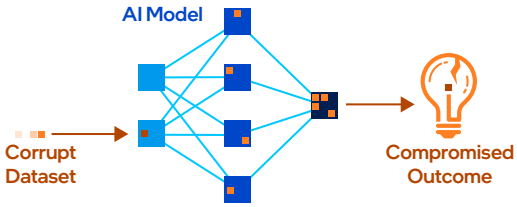


Mitigations:

Confidential Computing helps defend against model theft by ensuring the model is processed only in an unencrypted state within the protected memory space of an Intel-based TEE. This helps protect against direct memory scraping and other privileged attacks during inference. For AI models deployed at the edge (e.g., on medical devices, mobile endpoints, or industrial sensors), where physical and administrative controls are limited, running within a Confidential Computing enclave significantly elevates the security posture, approaching the assurance levels of secured data center environments. It also helps ensure that only attested and authorized code can access the model, making model extraction significantly more difficult and detectable.

Data Poisoning:

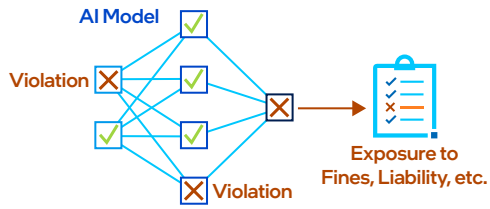
In the training phase of AI development, attackers can tamper with datasets—injecting mislabeled, adversarial, or biased samples to corrupt the learning process. This can lead to subtle but dangerous degradations in model behavior, such as misdiagnosis in healthcare models, discriminatory outputs in hiring tools, or exploitable behaviors in autonomous systems. Poisoning can occur through compromised data pipelines, third-party data sources, or malicious insiders—especially in federated or collaborative training environments.



Mitigations:
Confidential Computing is designed to ensure that training data is processed only within attested, hardware-isolated environments, shielding it from tampering by malware, rootkits, or even privileged insiders. This not only helps protect against zero-day exploits and system-level threats but also supports data provenance, enabling models to be trained only on verifiably unmodified inputs. When combined with encryption at rest and in transit, Confidential Computing enables full-lifecycle protection for sensitive training data, helping ensure integrity from data ingestion to model output.

Compliance Violations:

Industries such as healthcare, finance, and public services face mounting regulatory pressure to ensure responsible AI use, with frameworks like General Data Protection Regulation (GDPR), European Union (EU), Health Insurance Portability and Accountability Act (HIPAA), and others enforcing strict controls on how personal and sensitive data is processed, especially in AI systems. GenAI in particular introduces complex risks around data leakage, lack of transparency, and unexplainable decision-making. If a GenAI system inadvertently trains on or reveals protected data or operates without verifiable controls, organizations may be exposed to severe fines, legal liability, and reputational damage.



Mitigations:
Confidential Computing offers a powerful foundation to enhance compliance in AI by enabling attested execution environments, where only cryptographically verified code can access sensitive data or models. These attestations serve as machine-generated, tamper-proof evidence of data handling practices, increasingly satisfying requirements for data-in-use protection, software integrity, and provenance tracing. For Generative AI workloads—whether in training, fine-tuning, or inference—this helps ensure sensitive data is processed only within verified and isolated environments, supporting Zero Trust principles and aligning with evolving audit and governance frameworks.

Enabling Confidential AI at Scale

As AI platform solutions and services continue to expand across cloud, edge, and on-premises environments, there are many paths to scale. Cloud-first data and analytics providers are playing a pivotal role in driving AI adoption. Major hyperscalers are already integrating Confidential AI services for data clean rooms, multi-party federated learning, confidential inferencing, and more. OEM and ISV partners are integrating Confidential Computing into their platforms—whether through secure hardware, software frameworks, or data-centric services—to enable integrity and confidentiality in solutions that accelerate the broader growth and trusted deployment of AI across industries.

Public Cloud

Intel is working with its cloud partners to ensure that technologies such as Intel TDX and Intel TDX Connect become ubiquitous across all types of cloud environments—public, private, hybrid, and multi-cloud.

For example, Google's adoption of Intel AMX and Intel TDX to drive accelerated confidential workloads that can also leverage Nvidia GPUs is now available on Google Cloud. Alibaba Cloud shows how to integrate Haystack from Deepset into an Intel TDX instance to deliver a more secure RAG solution.

By embedding Confidential Computing at the core of AI cloud infrastructure, Intel aims to:

1. Build Trust: Enhance trust and security for sensitive AI workloads, allowing enterprises to innovate with confidence.
2. Streamline Integration: Simplify adoption by ensuring seamless integration of Intel TDX technologies into AI development and deployment pipelines.
3. Enable More Secure Migration: Accelerate AI workload migration to secure cloud environments, enabling broader enterprise adoption of cloud-native AI.
4. Expand Partnerships: Strengthen ecosystem partnerships by collaborating with ISVs, systems integrators, and service providers to optimize Confidential Computing solutions across the cloud AI stack.

Through this focus, Intel and its partners will help organizations unlock new AI capabilities while safeguarding data, models, and intellectual property—driving the next wave of secure, scalable AI innovation.

Advancing Confidential AI in the Enterprise

Leading OEMs are investing in building Confidential AI appliances, designed to meet the growing demand for high-performance, secure AI infrastructure in sectors such as healthcare, finance, and government. These appliances are purpose-built for enterprises that must process sensitive data within trusted, on-premises environments.

Intel's OEM partners are increasingly recognizing the value of Intel-based Confidential Computing technologies as a critical enhancement for large-scale deployments, particularly in highly regulated industries.

One example of innovation in this space is the development of PrivateGPT appliances—secure, enterprise-grade AI platforms that enable organizations to build, fine-tune, and deploy GenAI models on sensitive proprietary data without exposing that data outside their trusted environments. These Confidential AI solutions, powered by Intel technologies, enable organizations to unlock the full potential of AI while maintaining control over their most valuable information assets.

To support this growing opportunity, Intel is focused on two key areas:

1. Empowering OEM Confidential AI Stacks: Intel is enabling its OEM partners to integrate Intel's Confidential Computing technologies into their Confidential AI solutions, helping enterprises confidently adopt secure AI platforms tailored to their needs.
2. Driving CPU-Optimized GenAI Solutions: Intel is also collaborating with OEM partners to develop CPU-only GenAI appliances that deliver strong price-performance advantages, making secure generative AI accessible to small and medium-sized enterprises. Dell is leading the way, showing how Intel TDX helps protect AI workloads when running on Intel® Xeon® processor-based servers.

Through these strategic initiatives, Intel is advancing the next generation of secure, scalable AI solutions with its OEM partners, helping enterprises accelerate innovation while protecting what matters most in any environment.

Empowering the ISV Ecosystem

The growing ISV ecosystem plays a vital role in expanding the adoption of secure AI solutions across industries. Companies specializing in data security, confidential cloud services, and AI infrastructure are increasingly building on-premises and cloud-based Confidential AI solutions, driven by their customers' demands for greater data privacy and regulatory compliance.

Leading ISVs across the AI pipeline—from security providers to vector database companies, MLOps platforms, and model developers—are building the tools required for enterprises to develop, deploy, and manage GenAI capabilities securely. By engaging with these ecosystem participants, Intel is helping to enhance the security posture throughout the AI lifecycle and accelerate adoption of Confidential Computing technologies.

To support and strengthen this ecosystem, Intel is focused on:

1. Partner Engagement: Intel is collaborating early with key ISV partners by providing early access to Intel hardware, technical enablement, and architectural guidance to drive the integration of Intel Confidential Computing technologies into their AI offerings. Privatemode.ai enables enterprises to use large language models like Llama 3 securely by running the entire inference process—inputs, models, and outputs—inside a TEE. This Confidential Computing approach helps ensure that sensitive enterprise data, such as customer records or internal documents, remains encrypted and isolated, even while in use, shielding it from the cloud provider, infrastructure admins, and external threats. The model itself is also

increasingly protected from theft or tampering, and remote attestation provides cryptographic proof that only trusted, verified code is running. For enterprises, this means they can deploy GenAI solutions—such as secure copilots, document summarizers, or analytics tools—with increased confidence in their ability to maintain data confidentiality, meet regulatory requirements, and support data sovereignty across public and hybrid clouds.

2. Joint Innovation and Market Alignment: Intel is working closely with strategic partners to develop joint go-to-market initiatives, helping ensure that Confidential Computing solutions are well aligned to emerging industry needs and AI use cases. Intel is collaborating with Fortanix on solutions like *Armet AI*, which combines Intel TDX with Fortanix's secure data platform to help protect sensitive data and models in AI pipelines—enabling regulated industries to adopt Confidential AI with greater confidence and compliance.
3. Showcasing Industry-Specific Solutions: Intel is highlighting real-world examples and industry-specific deployments where Intel's Confidential Computing technologies help ISVs deliver differentiated, Secure and Trusted AI solutions to their customers. EQTY Lab integrates Intel TDX to build privacy-preserving AI infrastructure that enables enterprises to collaborate on sensitive data without sacrificing control or confidentiality. Their platform is designed around verifiable trust, ensuring that AI models operate only within attested environments, with strict enforcement of data usage policies. This approach empowers customers to adopt AI in high-stakes contexts where transparency, integrity, and data sovereignty are essential.

By fostering deep technical partnerships and empowering the next generation of AI-focused ISVs, Intel is positioning Confidential Computing as a critical enabler for secure, trusted AI innovation—helping customers unlock the full value of AI at scale while maintaining strict control over their sensitive data.

Confidential Computing: The Foundation for Secure AI Innovation

In a world where AI is rapidly becoming pervasive, the importance of Confidential Computing for improved security cannot be overstated. As organizations deploy AI across internal and external environments, protecting sensitive data at all stages—in transit, at rest, and in use—is no longer optional... it is essential. In many sensitive data scenarios, it could be the catalyst to unlock data silos that open huge digital transformations.

Intel's Confidential Computing portfolio provides comprehensive solutions to secure AI wherever it resides. It addresses the threats that GenAI introduces or exacerbates, giving organizations increased confidence to accelerate their AI journeys and unlock the full value of their proprietary data. This portfolio is powered by open software and a vibrant ecosystem of software, cloud, and device partners that extend its reach far beyond Intel's walls.

The future of Confidential Computing for AI workloads lies in its ability to provide businesses with the trusted foundation they need to innovate without compromise. With that strong foundation in place, Intel is advancing the Confidential Computing ecosystem to meet the evolving demands of AI—today and into the future.

Look for Intel-based Confidential Computing instances in Cloud Service Provider and server vendor offerings today. Contact Intel's many ecosystem partners for value-added solutions, or visit **intel.com/confidentialcomputing** for more information on how Intel is designing Confidential AI to be secure, scalable, and trusted by design.

Legal Notices and Disclaimer

i. IDC 2024 Global GenAI Technology Trends Report
ii. Confidential AI Survey, Q3 2024; IDC 2024 Global GenAI Technology Trends Report
iii. OWASP Top 10 for LLM Applications 2025; NIST Trustworthy and Responsible AI NIST AI 100-2e2025