DELL Technologies

5 Recommendations to Maximize GenAI Securely

1	2	3	4	5
Secure the layers of a GenAI system	**Utilize zero trust principles**	**Maintain governance and human oversight**	**Take advantage of GenAI security tools as they become available**	**Innovate with confidence**
Infrastructure	Never trust, always verify	Engage key stakeholders	Content	Aim for cybersecurity to facilitate the mission, not inhibit it
OS and Kubernetes	Least privilege access	Set policies for ethical and regulatory compliance, data management	Risk prediction	Let cybersecurity maturity build organizational confidence to innovate
GenAI Applications	System hardening	Monitor and enforce accountability	Knowledge and automation	
Data	Identity management	Training and education		
	Segmentation			
	Logging, monitoring, and auditing			

Generative AI technology promises transformative capabilities but comes with unique security challenges.

Generative AI is revolutionizing business like never before, driving innovation and offering unparalleled advantages that provide a competitive edge. While this technology has transformational potential, it also comes with its own set of security challenges.

Dell subject matter experts Steve Brodson, Services Product Manager and Eitan Lederman, Cybersecurity Consultant joined Chris Cicotte from the APEX and AI marketing team to address those concerns and discuss ways to Maximize GenAI Securely. Read on for a summary of the conversation and additional insights on the topic and watch the full discussion at dell.com/cybersecuritymonth.

> It's about training the people. People need to know how to use the GenAI system. What to do, but nevertheless also what not to do."
>
> ***Eitan Lederman***
> *Dell Cybersecurity Consultant*

Secure the layers of a GenAI system

While GenAI is a relatively new technology, most of the security protocols are the same established cybersecurity techniques that are used to secure other workloads.

Infrastructure – Focus on minimizing the attack surface:

- Vulnerability and penetration testing
- Patching
- Hardening
- Identity management, including strong passwords, multi-factor authentication (MFA)
- Monitoring and auditing
- Ensuring the third party supply chain is secure

OS, and Kubernetes - also a focus on attack surface reduction including:

- Vulnerability scanning
- Regular patching
- Updating Kubernetes components
- Limiting access control based on identity management, roles-based access (RBAC), and least privilege access
- Securing the control plane including the the API Server, secrets, kubelet and other components
- Using namespaces

GenAI Applications - Implement security actions targeted at the new attack surfaces created by GenAI:

- Identity management to address prompt injection, sensitive information disclosure, model theft, training data poisoning
- Data source validation to protect against training data poisoning, model bias
- Monitoring and auditing to identify and prevent model DOS, model theft, sensitive information disclosure, anomaly detection, forensics

Data - Incorporate strong data protection measures to secure the data in the language model and application:

- Air-gapped cyber vault
- Encryption
- Incident response plan
- Monitoring and auditing of training data and output

Ensure that data protection principles are applied to all data, including training inputs, model outputs, and any data involved in Retrieval Augmented Generation (RAG), if used. Additionally, ensure ongoing compliance with all applicable data protection regulations.

Utilize zero trust principles

The role of several zero trust principles like identity management, least privilege access, system hardening and patching have already been mentioned, indicating the value of zero trust principles in securing a GenAI workload. Zero trust architectures also require

DELL Technologies

continual logging, monitoring and auditing of network activity, which can prevent GenAI specific risks like result manipulation and data poisoning.

Moreover, zero trust also encourages micro-segmentation which reduces the impact of a breach. It also requires data encryption, both in transit and at rest, which is an important part of the overall data protection strategy.

While these are just some of the ways that zero trust can secure a GenAI workload, adopting zero trust principles should be considered a best practice.

Maintain governance and human oversight

Much of GenAI's value lies in automating tasks that humans would normally execute, but human governance is critical to ensuring security and proper functioning of the applications. A governance model will typically involve key stakeholders throughout the organization who set guidelines and requirements for ethical and regulatory compliance, data management policies and procedures, and ultimately enforce accountability.

Appropriate governance and oversight can help address problems like model overreliance, bias, result manipulation, sensitive information disclosure, and data poisoning.

Lederman also pointed out the importance of training, "It's about training the people. People need to know how to use the GenAI system — what to do, but nevertheless, also what not to do."

In addition to the risk posed by an organization's GenAI applications, there is also the proliferation of GenAI-enabled cyberattacks which often require human intervention. Examples include malicious actors using deepfakes to drive human behavior and phishing attacks made much more effective by more accurately mimicking a human's writing or speaking style. Ongoing training and education are some of the most effective ways to address these risks, again reinforcing the human element.

Take advantage of GenAI in security tools as they become available

While much of the focus is on risk, GenAI also has the potential to bolster security efforts. While these capabilities are in their infancy, they will offer benefits in three key areas:

- **Content:** Generating security policy, personalized training, data classification, and reporting
- **Prediction:** of risk and attack activity, suggestion of remediating actions
- **Knowledge:** Querying the environment (talking to the system), forensics, automation

GenAI's contribution to security tools could help maximize the capability of security teams, reduce costs, and enhance defenses. Take advantage of these solutions as they grow and mature.

Innovate with confidence

Most importantly, don't let security risks prevent you from leveraging potentially revolutionary technology. Efficiency, automation, cost reduction, problem solving, and driving creativity are just some of the ways that GenAI can transform business.

While GenAI requires robust and sometimes new cybersecurity measures, the goal should be to facilitate the organization's mission, not inhibit it. Developing the right cybersecurity strategy should give organizations the confidence to grow and innovate.

Learn how to address some of today's top cybersecurity challenges at **dell.com/cybersecuritymonth**

DELL Technologies