The Art of Threat Modeling

1 November 2023 - ID G00796906 - 47 min read

By: William Dupre

Initiatives:Security Technology and Infrastructure for Technical Professionals; Meet Daily Cybersecurity Needs

> Threat modeling is a form of risk assessment used to identify exposures and mitigations in a system. Security and risk management technical professionals need to use various techniques to instill a threat modeling mindset in their organizations.

Overview

Key Findings

- Organizations struggle to understand security threats that could impact their systems, resulting in limited quality or lack of controls.
- Threat modeling can be a time-intensive exercise because teams need to acquire advanced skills, especially when creating new models from scratch. This leads to a deficit of security requirements and practices within teams.
- Many organizations default to using popular threat modeling frameworks (e.g., STRIDE), failing to take advantage of the diversity of frameworks and techniques available.
- Threat modeling is typically seen as an application security tool or process, leading delivery teams to dismiss it as “the security team’s job.” Also, organizations are not taking advantage of the threat modeling approaches in other areas of digital risk.

Recommendations

To better understand threats, security and risk management technical professionals must:

- Use threat modeling frameworks as a basis for understanding threats within a system and to increase the quality and efficacy of implemented controls.
- Use a diversity of frameworks and techniques to get a more complete picture of threats against all digital systems throughout the organization.

- Provide teams with architectural standards and policies based upon threat modeling results.
- Enable teams to do threat modeling by providing automated solutions and a diversity of techniques.
- Instill a threat-conscious mindset by adopting threat modeling across the organization and having multiple roles involved in threat modeling processes.

Analysis

Threat modeling is an architecture-level process for reviewing a system design, enumerating threats and mitigations, validating controls and mapping out the attack surface of a system. This can be for an application, a network, a device, containers, or any system or element of software or hardware. It is a core part of the secure-by-design approach to system development.

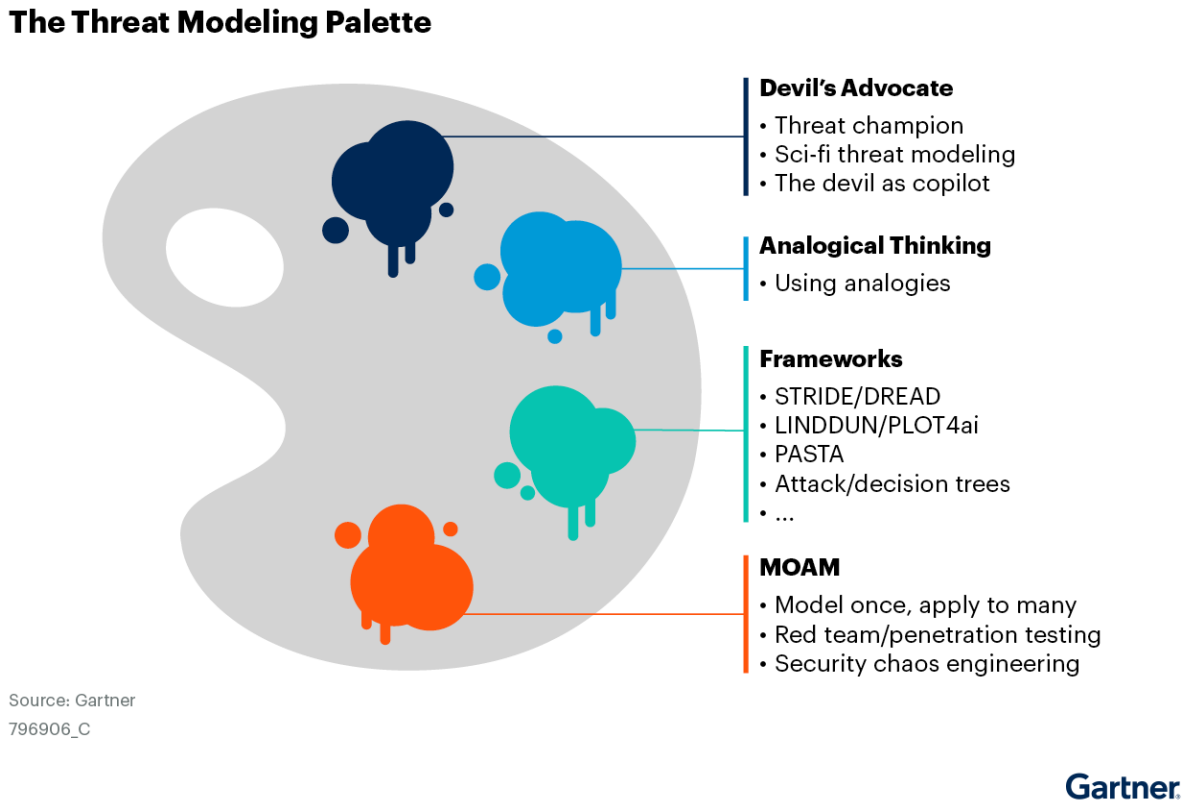
While the definition above expresses a formal approach to threat modeling, the goal of this research is to add a bit of art to the formal technique. A general starting point for threat modeling comes from the Threat Modeling Manifesto and centers around four key questions: [1]

- What are we working on?
- What could go wrong?
- What are we going to do about it?
- Did we do a good job?

The key lesson to be learned from these questions is that organizations need to go beyond adopting threat modeling tools and processes for cybersecurity and instill a certain mindset in their employees. In this way, employees will be more threat conscious and can use better judgment with respect to cyber activities. It is an important trait in a world where common threats persist because they consistently work and new threats emerge to adapt to new technologies and sociotechnical structures (for example, threats to and from generative AI [GenAI]). Just relying on formal processes and structures to address these threats will only add friction to threat discovery and anxiety to those needing to understand the threats.

This research will explore a palette of approaches to threat modeling, including both formal and informal approaches, to ensure teams feel like they are threat aware and not just working through a process. Figure 1 provides a sampling of the techniques that will be elaborated on in this research.

Figure 1: The Threat Modeling Palette



> **Software manufacturers should perform a risk assessment to identify and enumerate prevalent cyberthreats to critical systems, and then include protections in product blueprints that account for the evolving cyberthreat landscape.**
>
> *— Shifting the Balance of Cybersecurity Risk: Principles and Approaches for Security-by-Design and-Default*[2]

Use Common Frameworks as a Basis for Threat Modeling

Organizations should look to formal methodologies for threat modeling to form the basis of an overall program. Using the questions from the manifesto, a process for threat modeling using some common frameworks can be laid out. The following sections will provide guidance on addressing each of the questions.

What Are We Working On?

The first step in the process is to get a mental model of the system in question. One approach that is typically used when modeling threats (especially when using standard frameworks) is to create a conceptual view of the system with a data flow diagram (DFD). A DFD can be constructed in different ways, but the primary approach is to model a system using entities within a system, the relationships between those entities and the trust boundaries around the entities. The boundaries represent segments that divide different areas of trust. These could include the internet boundary, a network segmentation or anywhere a trust relationship changes. See Figure 2 for an example of a DFD.

What Could Go Wrong?

Using the DFD created above and a framework (or frameworks) like those discussed below, teams can now identify where threats could impact the system. Some organizations start modeling threats by doing this informally on a whiteboard or printed out diagram. Others may use threat cards illustrating specific threats (see below for card games). As the process matures, automated tooling will be used to identify threats within a system. See the sections below for different frameworks and vendors providing threat modeling tools.

To enhance this part of the process, security teams should bring in threat intelligence to better understand what threats may be feasible. The goal would be to use threat intelligence that is a combination of generic information and information specific to the organization and/or industry to improve the modeling process.

What Are We Going to Do About It?

The primary goal of threat modeling is to address any exposures in a system that could be exploited by threats. At this point, teams should be looking to identify controls that are needed to address those exposures. These could be technical controls (for example, a web application firewall or identity and access management controls), process controls or cultural changes. This process could be manual as well, but tools are available to identify controls.

Using a risk-based approach for this effort is important. Not all threats identified in a threat model have the same risk, and organizations should prioritize mitigations based upon risk appetite. This is a point where threat intelligence will be important as well. Security teams can contextualize threats based upon their own intelligence gathering efforts.

Within the context of risk and threat intelligence, teams can then determine what actions to take. Typical actions fall into four categories:

- **Eliminate or avoid:** If the threat doesn't exist, there is no risk from it. This may require changing the design or implementation of a system to remove the possibility of a threat. For example, if sensitive data that is being collected could expose the organization to an information disclosure threat, but provides no business or strategic value, then the system should be redesigned to not collect that kind of data.
- **Mitigate:** This approach requires the identification and addition of controls to detect and/or prevent attacks to an acceptable level. For example, if sensitive data with business value is at risk from an information disclosure threat, then identity and access management (IAM) or encryption controls may need to be put in place.
- **Accept:** All organizations have to deal with some level of risk — it is necessary to do business. Accepting the risk posed by a threat may be the most appropriate approach due to the complexity and/or cost of either eliminating or mitigating the threat. However, acceptance must be documented and done with full knowledge of stakeholders and reviewed periodically as the threat landscape evolves. Accepting a threat doesn't mean ignoring it.
- **Transfer:** This is the process where an organization shifts the consequences (for example, financial) of a risk to some other party. This can be done through an insurance policy so that an organization can recover costs incurred from an incident.

Beyond just applying mitigation measures to the specific system that has gone through threat modeling, organizations should consider those actions as possible architectural or organizational standards that could apply to systems with similar architectures and threat postures. This will provide development and security teams with standards to build secure systems without having to perform a repetitive modeling process.

Did We Do a Good Job?

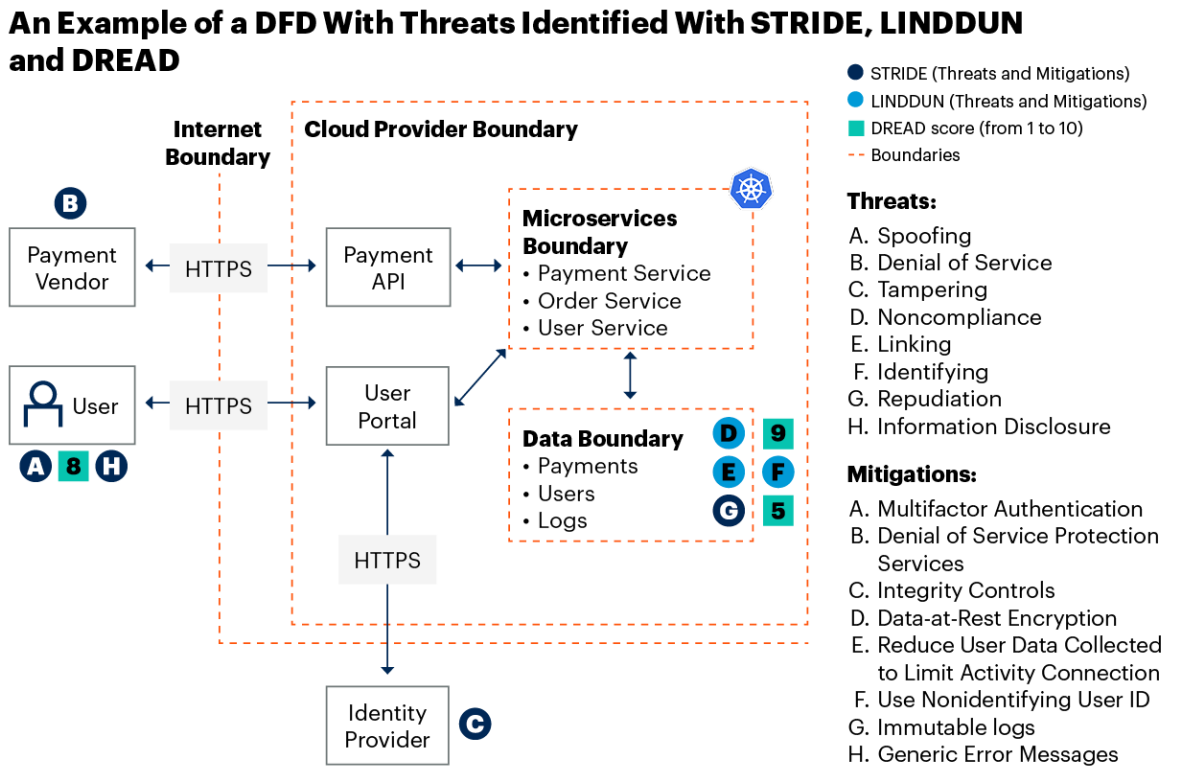
A key tenet of any process should be continuous improvement. Threat modeling is no exception. Teams should look over the process and determine what changes could be made to a specific threat model or the practice in general. For instance, they could determine that automation is needed to improve the efficiency of the threat modeling process.

When reviewing the process to determine areas of improvement, organizations should consider other aspects of systems that could be modeled. Examples could include illustrating data or system classifications in the model, providing visibility into regulatory concerns that impact threat risk, and modeling insider risk. It is also important to think about what other organizational roles could contribute to the understanding of threats and risk.

An important component of the overall process is documentation at every step. DFDs help to illustrate a system, but keeping track of threats and mitigations will be just as important. These artifacts can provide a current snapshot of system security, drive its evolution, and offer architectural guidance on threat mitigations. This can start with traditional systems such as wikis, file shares and ticket tracking systems, but threat modeling tools will make this effort more efficient, especially if they integrate with chat and tracking systems (for example, Salesforce's Slack, Microsoft Teams and Atlassian's Jira Software).

Figure 2 illustrates a DFD with threats identified using various frameworks. Those frameworks are discussed in the sections below.

Figure 2: An Example of a DFD With Threats Identified With STRIDE, LINDDUN and DREAD



Source: Gartner

Spoofing: Legitimate user's account taken over; Denial of Service: API endpoint taken offline by excessive amount of requests; Tampering: Identity provider data manipulated; Noncompliance: Financial data not protected appropriately; Linking: User linked to specific action; Identifying: User identity exposed; Repudiation: Logs manipulated to erase activities; Information Disclosure: Error message to user discloses internal application logic.

796906_C

Gartner

STRIDE

STRIDE is the most popular threat modeling framework. Gartner clients showing interest in threat modeling typically ask about (and use) STRIDE. The name is a mnemonic that stands for the following:

- **Spoofing:** An attacker attempts to impersonate an entity (for example, a user, a service) that interacts with some part of the system.
- **Tampering:** An attacker tries to modify data to manipulate some outcome.
- **Repudiation:** A user is able to deny actions, leading to a lack of attribution.
- **Information disclosure:** A system exposes information not intended to be released or that can be used for malicious purposes.

- **Denial of service:** Normal access to a system is restricted or prevented.
- **Elevation of privilege:** Permissions or authorizations not available to an entity are granted.

To model threats using STRIDE, teams start with a DFD, then use the threat mnemonic to determine where in the architecture a system could have exposures. For example, a client system that interacts with an API could be spoofed by an attacker causing a data leakage or breach. Another example would be to illustrate where an attacker could possibly elevate privileges to an administrative user if multifactor authentication is not in place. This process would continue across the DFD using all the STRIDE threats. The end goal would be to identify necessary security controls needed for the architecture.

The STRIDE card game called the Elevation of Privilege Game provides an easy way to identify various threats in a system. It can also be used to gamify the process and educate teams on security threats. [3]

DREAD

DREAD is a risk assessment model that can also be used to complement STRIDE (or other threat models) to help an organization understand the impacts of threats. “DREAD” is an acronym that describes five criteria to help assess threats. These criteria are:

- **Damage:** How bad would an attack be?
- **Reproducibility:** How easy is it to reproduce the attack?
- **Exploitability:** How much work does it take to launch the attack?
- **Affected users:** How many people will be impacted?
- **Discoverability:** How easy is it to discover the threat?

Each of these criteria can be scored with a rating from 1 to 10, with the average across all being the overall score for the threat. Figure 3 illustrates the DREAD rating and overall score of a hypothetical information disclosure threat.

Figure 3: The DREAD Results of a Hypothetical Information Disclosure Threat

The DREAD Results of a Hypothetical Information Disclosure Threat
Illustrative



Source: Gartner
796906_C

Gartner

DREAD should be used when a rough idea of risk is needed to help prioritization efforts. Organizations should consider using DREAD along with other models.

LINDDUN

Security and risk management technical professionals use the confidentiality, integrity and availability (CIA) triad to align controls to information security risks. While the confidentiality component of the triad is about protecting data against unauthorized access, it lacks the focus of protecting the rights of individual people in controlling their personal information. That is the domain of privacy.

Privacy regulations, such as the General Data Protection Regulation (GDPR) and the California Consumer Privacy Act (CCPA), are proliferating around the world. These are efforts by governments to provide citizens with better control and protection of their personal data. In general, these regulations have the privacy protection goals of unlinkability (personal data cannot be linked across domains), intervenability (individuals have control over the processing of their personal data) and transparency (data processing of personal information can be understood and reconstructed at any time). Security professionals should see these goals as a complement to the CIA triad and align their controls to them when data privacy is a concern.

Organizations that have concerns about data privacy should include a more focused threat modeling approach. One such framework is LINDDUN which provides a catalog of privacy threats to enable the investigation of a wide range of design issues that could impact privacy. The acronym "LINDDUN" represents the following privacy threat types:

- **Linking:** ability to associate data or actions to an individual or group.
- **Identifying:** learning the identity of an individual.
- **Nonrepudiation:** being able to attribute a claim to an individual.
- **Detecting:** deducing the involvement of an individual by observing.
- **Data disclosure:** excessively collecting, storing, processing or sharing personal data.
- **Unawareness:** insufficiently informing, involving or empowering individuals in the processing of personal data.
- **Noncompliance:** deviation from security and data management best practices, standards and legislation.

The methodology consists of modeling a system (using a DFD), eliciting threats (using a mapping table and threat trees) and managing threats (prioritizing and mitigating identified exposures).

When modeling threats with LINDDUN, it is important to consider the roles involved in the process. Legal and compliance teams should take part in the effort to better inform the organization of the legal and regulatory aspects of privacy. Security and IT teams should also be involved to make sure the necessary technical protections can be put in place.

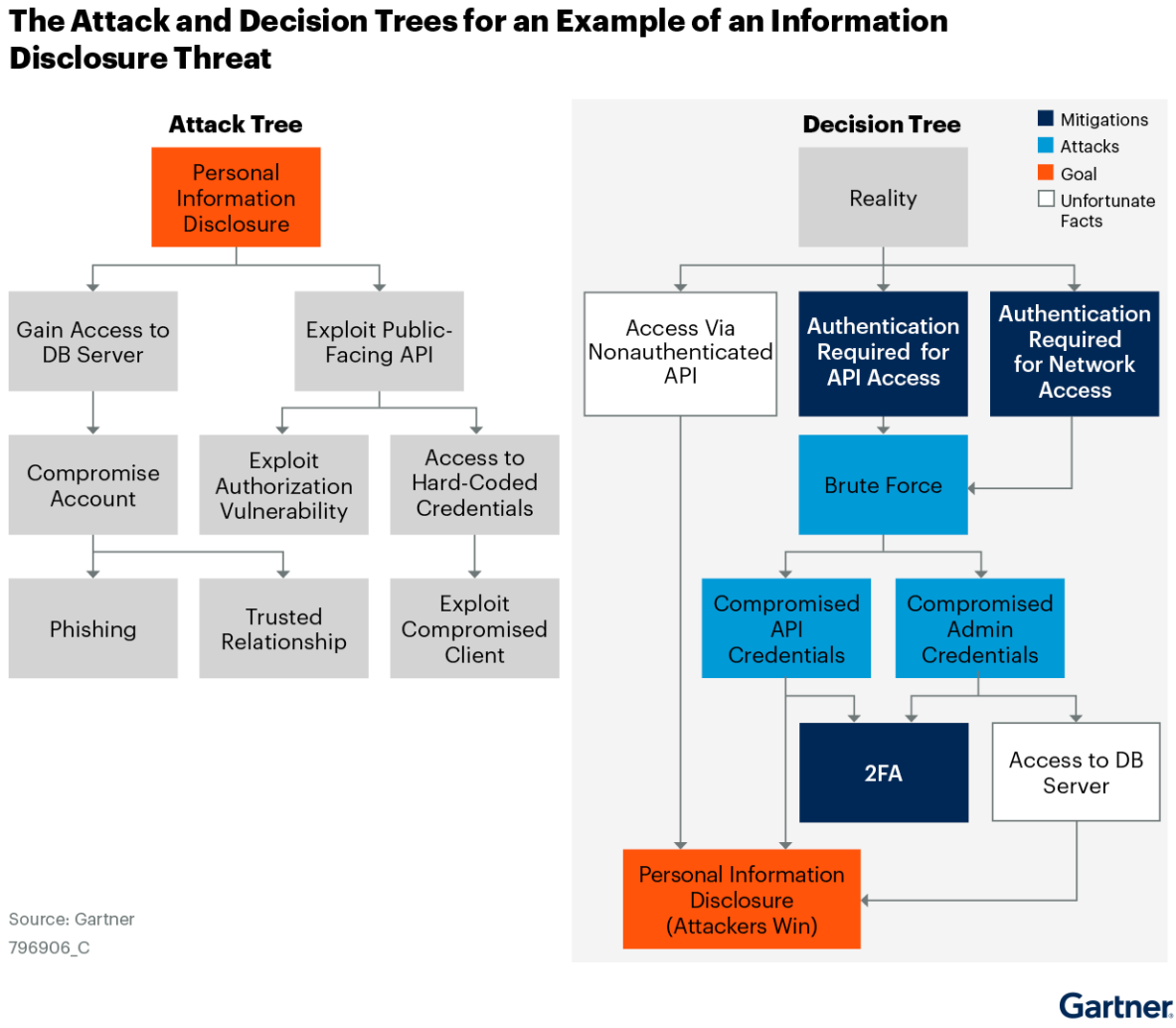
Attack Trees and Decision Trees

Attack trees are an attacker-centric threat modeling technique that allows teams to model how an attack might unfold. The attack scenario is modeled as a tree structure with the root node being the goal of the attack and leaf nodes being approaches for achieving that goal. Unlike other threat modeling frameworks, attack trees do not use a DFD.

An alternative to attack trees is a decision tree which models the actions an attacker might take at each stage of an attack and what a system can do to counter the attacker. [4] This approach can help teams understand the attacker mindset and decision-making process, along with the return on investment (ROI) of their attack.

Attack and decision trees can be used as a stand-alone approach or as a further analysis complementary to other types of threat modeling techniques. For instance, once a threat is defined using STRIDE, an attack or decision tree can be created to illustrate the attack sequence of that possible threat. The process could be further enhanced with adversary tactics and techniques as defined by the MITRE ATT&CK framework or the Lockheed Martin Cyber Kill Chain framework. See Figure 4 for examples of attack and decision trees.

Figure 4: The Attack and Decision Trees for an Example of an Information Disclosure Threat



Others

The above modeling frameworks and techniques are popular, but don't constitute an exhaustive list. Table 1 lists the frameworks mentioned above along with others that are available.

Table 1: Threat modeling frameworks

(Enlarged table in Appendix)

Framework	Description	Focus
Attack trees/decision trees	Model attacks from the perspective of an attacker.	Attacker
DREAD	Risk assessment based upon damage, reproducibility, exploitability, affected users and discoverability.	Risk impact
LINDDUN/PLOT4ai	Catalogs of privacy threats for information and artificial intelligence/machine learning (AI/ML) systems.	Privacy
Operationally Critical Threat, Asset, and Vulnerability Evaluation (OCTAVE)	Used to identify and assess risks to IT assets. It assesses organizational risks, not technological risks.	Asset risk
Process for Attack Simulation and Threat Analysis (PASTA)	A seven-step, risk-centric modeling framework that focuses on the business impact of threats.	Business risk
STRIDE	Identifies threats based upon spoofing, tampering, repudiation, information disclosure, denial of service and elevation of privilege.	Threats
Trike	Used for security auditing processes to identify whether an assigned level of risk is "acceptable" for an asset.	Audit risk
Visual, Agile, Simple Threat (VAST) modeling	Covers the entire software development life cycle (SDLC) across an organization using the three pillars of scalability: automation, integration with SDLC tools and collaboration with cross-functional stakeholders.	SDLC risk

Source: Gartner (November 2023)

When to Conduct Threat Modeling With Frameworks and With Whom

Organizations should use these frameworks at selective points within risk assessment processes to get a broad and consistent understanding of threats. These points should align with the different phases of system evolution, such as:

- **Initial system design:** When a new system is being designed, performing threat modeling in the early stages, before coding starts, will be important in verifying the design and driving architecture standards and requirements for development.
- **Legacy system reviews:** In this context, a legacy system is one that was developed and put into operation before threat modeling was conducted. Even in this case, organizations should look to create an architecture view of the system and perform threat modeling on the system. It may be more difficult or costly to apply controls in this case, but it is crucial that threats to legacy systems are understood and mitigated where possible. This effort should not be considered an all-or-nothing effort. Modeling an entire production system may be a herculean task, so it may only be feasible to model certain aspects of a system or the entire system but at a more reasonable abstraction level.
- **Tactical architectural updates:** No system stays static over time; they all evolve. Threat modeling done at one point will lose its relevance quickly as new development proceeds. However, this does not mean that teams must perform threat modeling for every small change to a system. But one threat modeling process should be done when significant changes are being made. What is significant? There is an art to that as well. Examples could include moving from containers to serverless deployments, implementing changes to regulations (the GDPR, the CCPA, the Health Insurance Portability and Accountability Act [HIPAA], etc.), changing from web to mobile clients, using third-party APIs, moving from on-premises to SaaS solutions for corporate applications, and installing Internet of Things (IoT) devices.
- **Strategic IT or business changes:** Digital transformation is constantly occurring and, sometimes, organizations need to make significant decisions on their IT strategies. In these cases, threat modeling should be a key aspect of a digital transformation risk assessment. For instance, if a strategic decision has been made to move all applications from an on-premises data center to cloud-native deployments, then planning should include threat modeling. Also, threat modeling should be done when exploring new business opportunities, like entering a new regional market or during a merger and acquisition process.
- **Incident postmortems:** After a security incident, security teams, along with relevant stakeholders, should review the system's threat model to determine if new controls are needed. If a threat model doesn't exist for a system, this may be the perfect opportunity to rectify that deficiency.

It is important that a diversity of roles within the organization are involved in these exercises so that a complete picture of the threat landscape can be illustrated. For instance, business roles help provide context to what is being modeled, while security roles provide guidance on how a threat could unfold. Architects and developers will also provide insights into components of the application and infrastructure.

Getting a diversity of roles to participate in this time-consuming effort can be a challenge, though. This is where security teams need to express the benefits of such approaches by having the different roles understand how their specific knowledge helps put threats in context and provide guidance on mitigation options. What's in it for them are more secure products, less rework on bolt-on security and a better secure-by-design approach to system development.

If an organization has a security champions program, then it should look to those roles to help with the threat modeling process.

Challenges and Benefits of Using Threat Modeling Frameworks

Using threat modeling frameworks should be a core part of a risk assessment process. Organizations should determine which framework is appropriate for a particular system based upon what they are trying to target — risk, understanding threats and/or privacy.

One challenge of using these frameworks is that they can be a labor-intensive exercise for many organizations. However, such activities can provide a shared learning experience for teams. There are also ways to gamify these efforts to address the possible boredom that could arise. For example, the Elevation of Privilege Game, LINDDUN GO, and OWASP Cornucopia are card games that help create a more collaborative and enjoyable environment for teams.

Even with gamification, however, manually performing threat modeling using these frameworks can be an expensive endeavor. To reduce the costs, organizations can look to tools that can automate the effort. Vendors in this space include IriusRisk, Security Compass and ThreatModeler. Other solutions include Computer Aided Integration of Requirements and Information Security (CAIRIS), Microsoft Threat Modeling Tool, OWASP Threat Dragon and Threagile.

In general, identifying threats to a system helps organizations plan and implement security controls, and enable a more threat-conscious mindset. The additional benefits of using the above frameworks include:

- **Structuring the practice:** Teams have a place to start with threat modeling using a common framework and library of threats to guide the practice. This helps improve the consistency and repeatability of threat modeling across the organization.
- **Providing enterprise architecture standards:** Mitigations and controls identified during the threat modeling process should provide guidance for other systems and architectures in the organization. For example, if an application is vulnerable to a spoofing attack because it lacks multifactor authentication (MFA), then it might be appropriate to make MFA the standard for all applications with similar characteristics. This will provide teams with standards for a secure-by-design or "paved path" approach to application development.
- **Experimenting with prospective architectures:** A threat model allows teams to test ideas regarding architectural approaches to development, whether it is a new system design or an addition to an existing design. For example, if a development team is looking to move from containers to serverless deployments, a threat model can give them insights into the implications and costs of such a change.
- **Improving incident response:** A threat model can provide valuable information on where a system may be vulnerable. Incident response (IR) teams can use this information as part of an IR exercise to test assumptions about the system. They can also look to the threat model after an incident to get a better understanding of what may have gone wrong. Some threat modeling frameworks may do better at this than others (for example, decision trees).
- **Supporting compliance:** As more scrutiny of cybersecurity comes from regulatory agencies, threat modeling will become an important component of compliance practices. In April of 2023, the U.S. Cybersecurity and Infrastructure Security Agency (CISA), in partnership with several other agencies and governments, issued guidelines for secure-by-design practices. The agencies "recommend manufacturers use a tailored threat model during the product development stage to address all potential threats to a system and account for each system's deployment process." [2] While this is currently just guidance, more forceful approaches could be coming.

- **Contextualizing vulnerabilities:** Many organizations struggle to prioritize vulnerabilities being reported from their various security tools (for example, application security testing and vulnerability management tools). A threat model can put these vulnerabilities in context. For example, if an application with a high-severity vulnerability is being deployed into an environment shown to be vulnerable to threats, then that vulnerability may need higher priority than one being deployed to a more secure location. Teams need a mental model of threats in a system to better triage vulnerabilities and plan remediation.

Use a Diversity of Techniques to Understand Threats

Using frameworks to perform threat modeling is an important component of an organization's risk assessment process. The frameworks should be used, as a manual effort or along with automated solutions, to understand threats. However, instilling a threat-conscious mindset into the organization will take a diversity of techniques. Some of the techniques are informal, and those, listed below, get to the heart of the art of threat modeling.

The Devil's Advocate

Organizations need to break out of their echo chambers and have their assumptions challenged. Complacency can breed insecurity. What is needed is for organizations to have people and/or processes in place to arouse them from the sleep of the status quo.

The devil's advocate was someone identified by the Catholic Church to argue the case against canonization (sainthood) of a candidate. They would take the contrary position from the "promoter of the cause" who was tasked with making the argument in favor of canonization. Organizations can use this concept in various ways to challenge their assumptions and biases toward systems and threats. The sections below will explore some approaches to the devil's advocate idea.

Threat Champion

Many organizations employ the use of security champions to be the voice of security on distributed teams. Security champions are not security experts, but staff with enough knowledge and judgment to raise the concerns of security throughout the organization.

A threat champion, by contrast, would be the foil to security by thinking like an attacker or taking positions contrary to convenient system assumptions. This role could be played by security champions, pair programmers or anyone with a skeptical (possibly cynical) mindset. A threat champion could support teams with activities, such as:

- **Writing abuser stories:** Product teams typically write user stories to explain needed functionality for a software feature. These stories are written from the perspective of the end user, are narrowly scoped and include specific requirements for the functionality (including nonfunctional requirements). A threat champion could write a corollary to a user story from the perspective of an attacker or malicious user. The abuser story puts focus on how the application could be abused if necessary controls are not put in place. See Table 2 for some examples of user and abuser stories.

- **Playing fortunately/unfortunately:** The art of gaming can provide organizations with skeptical skills needed to understand threats. A threat champion can take a counterapproach to a positive scenario. For example, a team member might say, "Fortunately, we have a web application firewall (WAF) in place to protect our APIs." The threat champion could then counter with "unfortunately, our WAF is not so adept at protecting us against API-style abuses and attacks." This exercise could lead to a discussion on more appropriate controls or trade-offs.

- **Being Socratic:** A simple way to instill a threat-conscious mindset is to ask probing questions about a system or design. The basic question comes from the threat modeling manifesto: What could go wrong? But this process can be more elaborate and specific for the context. See Table 2 for some examples.

Table 2: Examples of User and Abuser Stories

User Story	*Abuser Story*	*Socratic Question*
As a payment app user, I want to transfer money from my account to the account of another app user.	As a malicious user, I want to transfer money from a legitimate user's account to my account without their consent.	What do you really know about your users? Do all of them have good intentions? Is authentication enough?
As a commenter, I can add comments to a blog post.	As a malicious commenter, I can add JavaScript to a blog post comment that executes a malicious function in a blog reader's browser.	Are you sure you want just anyone to add comments? Should we be maintaining the comments section or would it be better to use a third-party service?
As an employee, I can enter expense information into the reporting system to get reimbursed for travel expenses.	As a malicious employee, I can falsify expense records to get money transferred to my account.	What do you really know about our employees? Do all of them have good intentions?

Source: Gartner (November 2023)

Finding threat champions is similar to finding security champions. Organizations should start with those people who show some interest in security already, and build on the curiosity with mentoring and training. This could include having them participate in formalized threat modeling or tabletop exercises. Also, providing incentives for threat identification will be important as well. Whether it is some form of recognition or providing a token of appreciation, threat champions should be encouraged, not treated as pariahs trying to cause problems.

Science Fiction Threat Modeling

The Gartner Futures Lab is using the idea of science fiction storytelling to enable "technology innovation leaders to use exciting and intriguing stories to explain technological breakthroughs and imagine how they will affect society and industry" (see Quick Answer: How Can I Use Science Fiction Storytelling to Communicate a Vision of the Future?). Similarly, organizations can use this approach to explore security scenarios that could be plausible in the future. This technique is called science fiction or sci-fi threat modeling.

As an illustration, assume that an attacker had collected data on key personnel in an organization, including their personal data and social connections. Could the attacker, using social engineering and technological methods to disrupt the personal lives of those employees, impact the availability and/or morale of the workforce? Some employees may need time off because of credit or financial problems caused by identity theft. Other employees may have to deal with healthcare problems due to the tampering of health records or medical IoT devices, and some others may be put into an ethical bind due to blackmail or bribery. How would reduced availability and morale, along with insider risk, affect an organization's ability to respond to a technical incident like a ransomware or distributed denial of service (DDoS) attack? A focused attack on key people in the organization could cause major disruptions to an incident response. This "denial of workforce" threat could become more feasible as the costs of data collection and AI/ML technologies continue to fall.

One approach to sci-fi threat modeling comes from the graphic novels developed by Arizona State University's Threatcasting Lab to speculate on future cyberwarfare. [5] While this approach is not feasible for most organizations, the exercise in telling such a story could be enlightening. For example, organizations could play out a speculative scenario in a tabletop exercise to explore the impacts of planned business or IT strategy changes. See Maverick Stories: Brave New Worlds for further details.

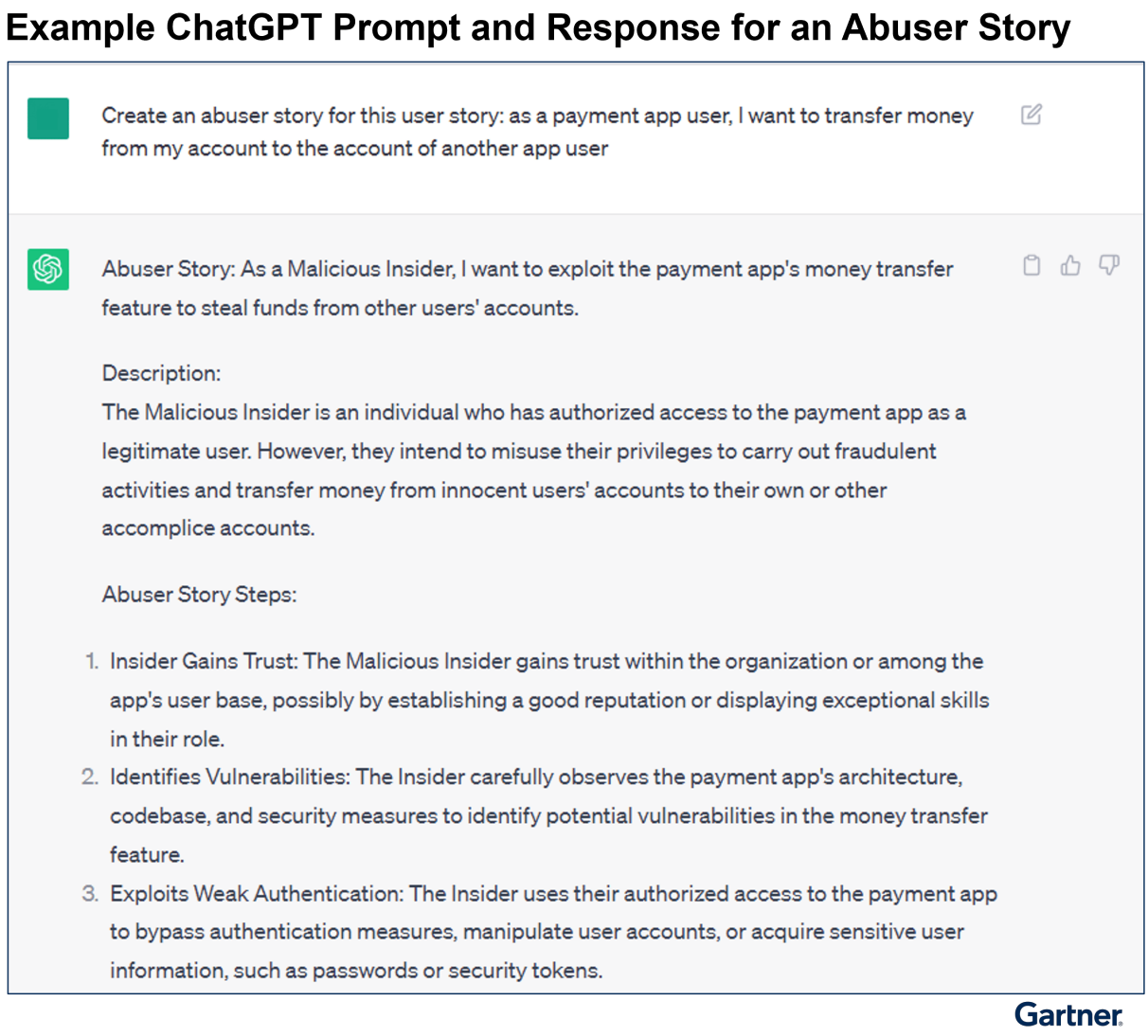
The Devil as Copilot

As AI capabilities improve, especially GenAI and large language models (LLMs), these systems can start to play the role of the devil's advocate. Threat intelligence and security operations systems will become more capable and provide organizations with the ability to improve incident response processes and threat modeling activities (see Quick Answer: How Can Security Operations Teams Leverage ChatGPT?). Organizations could use these capabilities to write abuser stories or sci-fi threat modeling scenarios.

It is key for organizations to improve their prompt engineering capabilities. A prompt is a description of the task that a GenAI model should perform, often in natural language. The following are some examples of prompts:

- Create an abuser story for this user story: "As a payment app user, I want to transfer money from my account to the account of another app user." See Figure 5 for ChatGPT's response to this prompt.
- What threats should be considered when implementing a set of REST-based APIs for a financial services organization?
- Speculate about cyberthreats to the U.S. electric grid in the year 2030.

Figure 5: Example ChatGPT Prompt and Response for an Abuser Story



As of this writing, AI solutions have been shown to have issues such as “hallucinations” where AI systems generate false information. Therefore, although AI tools can be very helpful in providing guidance and creative suggestions, security and risk management technical professionals should use them with caution and always validate their output.

Analogical Thinking

Having a threat-conscious mindset is more than just thinking about threats as they are. Sometimes it is important to be able to draw connections between something that is familiar and something that is alien to our understanding. Such connections can help us better understand new and complex issues and enable us to communicate such issues. This is the idea of analogical thinking or thinking with analogies. Security and risk management technical professionals should use such techniques to better understand and communicate threats.

> **Analogical thinking: thinking characterized by extrapolations from the familiar to the unfamiliar, rather than by the use of formal logic or consecutive reasoning. It is particularly important in problem solving and learning, in which known similarities between aspects of certain entities are used to make assumptions about other aspects or entities.**
>
> — *American Psychological Association*[6]

In his 2023 book “Threats: What Every Engineer Should Learn From Star Wars,” Adam Shostack uses this idea to explain the STRIDE threat model framework with themes from “Star Wars.”

For example (not from Shostack’s book), in the movie “Star Wars: A New Hope,” after Luke Skywalker and his uncle purchase R2-D2 and C-3PO, Luke is in his workshop cleaning the droids. R2-D2 shows a short hologram of Princess Leia that Luke wants to see more of. R2-D2 claims that a restraining bolt on him is preventing him from showing more of the hologram. Luke is convinced to take the bolt off, but R2-D2 refuses to show more of the clip and later escapes to search for Obi-Wan Kenobi.

This scene illustrates a social engineering threat. R2-D2 uses the pretext of an interesting hologram of a mysterious woman to get Luke to do something he shouldn't do. An analogy like this can help people understand and communicate a social engineering threat.

Model Once, Apply to Many (MOAM)

Organizations typically have hundreds, if not thousands, of applications and systems in their environment. When looking to conduct threat modeling on these systems, there is an obvious sense of anxiety as to how to approach such an effort. It is not feasible to attempt to take these systems one by one and perform a time-consuming threat modeling process. However, what must be considered is that not all of these systems or applications are unique in every dimension. It is possible that lessons learned from one threat model can be applied to multiple systems or applications. This is the idea of "model once, apply to many (MOAM)."

An organization that has over a hundred microservices deployed to a Kubernetes cluster with an API gateway in front to mediate external API connections may not need to perform threat modeling on every service or API. For example, one threat model created with STRIDE concerning the basic architecture could be enough to understand exposures across the service and API landscape. One caveat may be if certain APIs have more sensitive information and require unique threat visibility or if there are multiple client technologies using the APIs (mobile apps, single-page applications, IoT devices).

The crux of the MOAM approach is really about learning lessons from security activities and applying them to similar areas. But organizations shouldn't just limit this idea to threat modeling: They can learn lessons about threats from other security activities as well.

Red Team and Penetration Testing

Both red team and penetration testing exercises are carried out from the point of view of an adversary against an organization's defenses. The difference is that penetration tests are wide in scope and try to understand exposures broadly, while red team exercises have certain goals, such as achieving access to a sensitive system. Both approaches can help organizations better understand threats.

These exercises typically have the following phases:

- **Planning phase:** During the planning phase of a red team or penetration testing exercise, organizations should explore the threats that could impact their systems. This enables better planning for the exercise and allows the organization to form a mental model of what the real threats are. Using an existing threat model could also help guide this planning by informing the creation of incident response playbooks. The planning should be done in coordination with the team responsible for performing the exercise (whether internal or external). If performing a penetration test, then threat models can help guide testers in their efforts by providing scope and focus.

- **Exercise time:** During a red team exercise, organizations will have to react to what is happening in the environment, typically with a blue team (possibly a purple team as well). During this time, blue teams should be considering what real threats could occur based upon what they know about their organization and existing threat model. They should also keep track of activities in order to provide input into the postmortem effort.

- **Postmortem:** Once the exercise is complete, organizations should use the results to adjust their mental model regarding their security defenses. This will also inform them about what threats could really exist and how they should adjust their incident response processes and playbooks.

A key lesson that should be learned from these exercises is that technical systems shouldn't be the only focus when looking to mitigate threats. Organizations must understand how their people respond as well. The real exposures that a threat may exploit are sociotechnical, which cover people and process, not just technical holes.

Security Chaos Engineering

Chaos engineering is the use of experimental and potentially destructive failure testing or fault injection to uncover vulnerabilities and weaknesses within a distributed complex system. Such an approach, through experimentation, can test the resilience of a system and discover where it could fail before it actually fails. Security chaos engineering (SCE) applies this approach to the security resilience of a system. [7]

A key principle of SCE is that failure is inevitable and should be considered an opportunity for learning and system evolution. This contrasts with the traditional approach of treating failure as something to be feared and avoided at all costs. The former focuses on outcomes, like making the system "antifragile," and psychological safety while the latter focuses on "security theater" — performing activities to promote the idea of "doing something" about security.

> **Woods' Theorem: As the complexity of a system increases, the accuracy of any single agent's own model of that system decreases rapidly.**
>
> *— STELLA: Report from the SNAFUcatchers Workshop on Coping With Complexity*[8]

SCE enables organizations to adopt a resilience-based approach using two tiers: evaluation and experimentation. The goal of evaluation is to characterize a system's critical functions and interactions, along with its safety boundaries (similar to traditional threat modeling). Artifacts from the evaluation tier include a diagram of system data flows, documentation of critical functions and a diagram of failure behaviors. Decision trees can be used to perform such an evaluation.

Experimentation, on the other hand, is about exposing system behavior by running stress tests against the actual system. These "chaos experiments" help organizations understand the impacts on critical functionality when they are exposed to an adverse scenario or condition. Such experiments challenge our assumptions about a system by providing actual information about how a system behaves, not how we think it behaves.

An SCE experiment can test cloud infrastructure security controls. For example, a security assumption may be that deploying containers with severe vulnerabilities in a Kubernetes cluster should be blocked (this assumption may come from an evaluation phase or some other threat model). Then an experiment can be carried out by deploying a container with known vulnerabilities in the cluster to see how the system reacts. Will the system allow the container to be deployed? Will any alerts be triggered to indicate an attempt has been made?

There are two ways SCE can apply to threat modeling:

- Using evaluation to understand where a system may be vulnerable to threats. This gives teams more insights into what threats may succeed.
- Using experimentation to validate the assumptions identified during the evaluation phase or in a traditional threat modeling activity (for example, while using STRIDE).

Organizations should use SCE to better understand and model threats. Such techniques can be very valuable for validating or invalidating assumptions about threats, and enable a system to adapt to stressors.

Threats Beyond Vulnerabilities

There is a tendency in the security industry to view security problems as technology problems, whether it's unpatched vulnerabilities or misconfigured systems. Yet, threat actors can take advantage of systems even if those systems are free from technical exposures. The primary concern here is the threat of abuse of functionality — how legitimate capabilities of a system can be exploited for illicit behavior. [9]

An example of this problem is when malicious actors use their bot armies to take advantage of a system to commit fraud. They can sign up for or take over a large number of accounts and perform legitimate actions, such as buying popular merchandise, signing up for new user benefits and using financial services. The end goal, however, is to sell the merchandise illegally (for example, ticket scalping) or to launder money. [10,11,12]

While there are technical controls that can be put in place to address such threats, the primary exposure is around the business decision on providing the functionality. This illustrates the importance of conducting threat modeling and thinking differently about threats as early and as often as possible.

Use Threat Modeling in Other Areas of Digital Risk

Organizations have typically used threat modeling to understand threats to an application or API. However, the Threat Modeling Manifesto is not solely focused on application security. The four questions that it includes aim to instill a threat-conscious mindset in practitioners. Regardless of what is being designed or implemented (an application, a cyber-physical system, a cloud strategy), organizations should look to threat modeling to better understand risk.

Cloud-Based Enterprise Applications

Many organizations are moving their corporate applications (for example, ERP, HR, payroll) to cloud-based solutions. While this approach does provide some security benefits as cloud providers take on some of the security responsibilities, organizations still must align with the shared responsibility model inherent in cloud usage.

With the four questions in mind, organizations should first understand what they are doing when moving to cloud-based solutions. Usually, they need to move data and services, many of which could be highly sensitive, to clouds. This will also involve new processes for employees in various job functions (HR, application development, finance, etc.).

With a threat-conscious mindset, an organization should ask: “What could go wrong?” Answers to this question may include:

- Brute force breaches due to misconfigured access control.
- Data integrity compromised due to improper role-based access control (failure to implement the principle of least privilege).
- Data exposure from provider breaches due to insufficient due diligence performed on service providers.
- Account takeovers due to social engineering, leading to information disclosure or data tampering.
- Data disclosure, data tampering or privacy violation due to insider risk.

Organizations can now think about what they are going to do about it. Answers could range from stronger access control policies (for example, multifactor authentication), improved due diligence during service procurement processes and implementation of security service edge technologies (see Adopt Security Service Edge (SSE) to Replace Stand-Alone SWG, CASB and ZTNA Products).

Cyber-Physical Systems

Cyber-physical systems (CPS) are systems engineered to orchestrate sensing, computation, control, networking and analytics to interact with the physical world (including humans). When secured, they enable safe, real-time, reliable, resilient and adaptable performance. The increasing number and interconnectedness of these systems, including those deployed using wireless technology, pose serious challenges to security professionals and require a different approach to threat analysis.

Examples of cyber-physical systems include:

- Medical monitoring devices
- Autonomous automotive and avionics systems

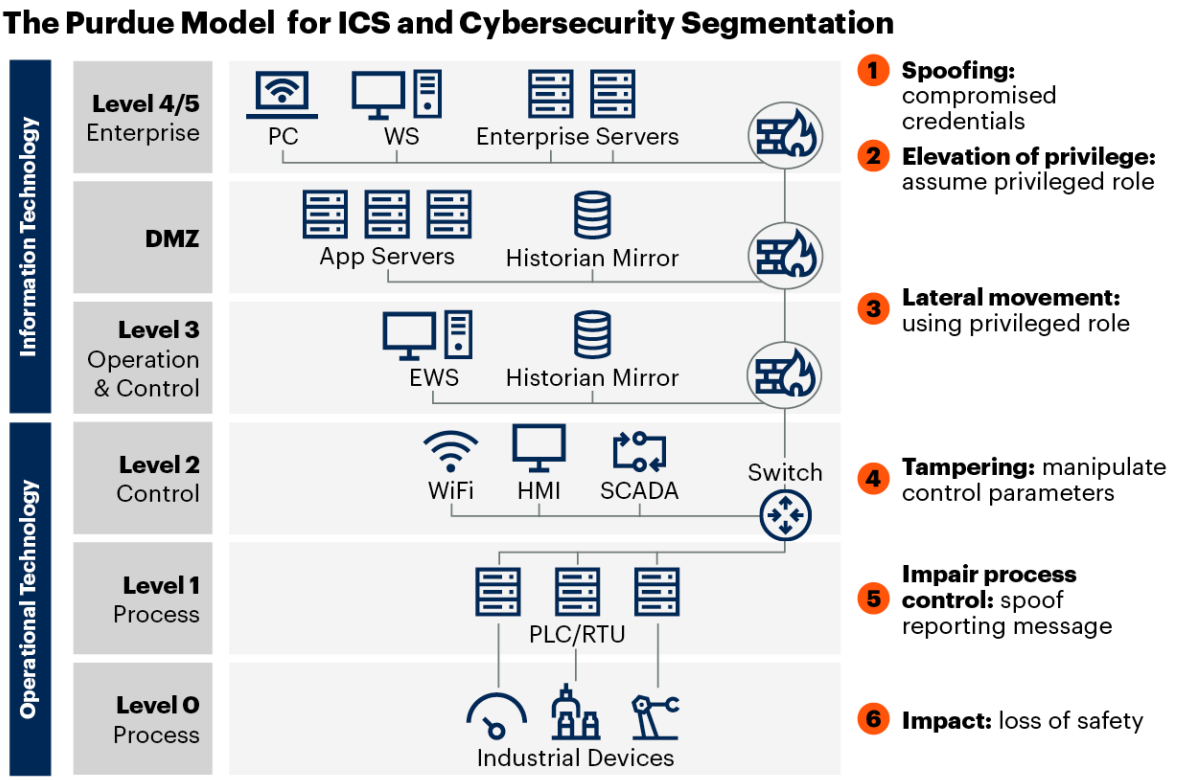
- Industrial control systems (ICS), including supervisory control and data acquisition (SCADA) systems and programmable logic controllers (PLCs)
- Internet of Things (IoT) and industrial IoT (IIoT)
- GPS devices

Most organizations still think about security only through the lenses of IT and data protection. However, to support resilience and growth as the business moves beyond IT environments, security professionals must expand their threat horizons. They need to understand the threats across the cyber and physical planes under a CPS security umbrella, and also understand that human safety could be a consequence of unsecured systems. They must understand that a successful CPS attack is different from a typical cyberattack on an IT system. Operational technology (OT) environments are each more unique than most IT environments, which use similar architectures and products. The attacker needs specialized knowledge and skills to understand "signal processing, control principles, the physics of process behavior, the mechanics and failure conditions of the equipment, etc." [13] In the end, "the attacker needs to interact with the control system to achieve the desired outcome in the physical world." [13] This requires different approaches to threat modeling.

A high-level approach to conducting threat modeling on CPS is to use the Purdue model for ICS security. This model can help security professionals understand how IT and OT systems should be segmented. It can be coupled with the MITRE ATT&CK framework for ICS to model threat tactics and techniques. Attack or decision trees could also be used to model attacker decision making across the Purdue model.

More traditional threat models could also be used in this approach. For example, STRIDE could be used to model the initial access to the enterprise IT system. From there, lateral movements to lower levels could be modeled using techniques from MITRE. Figure 6 provides a simplified threat model using the Purdue model.

Figure 6: The Purdue Model for ICS and Cybersecurity Segmentation



DMZ = demilitarized zone; EWS = engineering workstation; HMI = human-machine interface; PLC = programmable logic controller; RTU = remote terminal unit; SCADA = supervisory control and data acquisition; WS = workstation
Source: Gartner
796906_C

While the Purdue model can provide a high-level view of how an attacker may move through an organization's network(s), it lacks a focused view of assets in the field. Those assets may have physical characteristics that don't react cleanly to changes in the digital environment. At some point, physics, not information technology, becomes the controlling factor. For instance, how a physical cooling system reacts to a change in the digital environment may be highly time- and environment-dependent and not fully understood by the attacker. Such processes may enter an uncontrollable state that significantly impacts the goals of the attacker (unless, of course, they are strictly looking to inflict chaos). Also, some wireless devices, such as GPS devices, can be subject to jamming and spoofing. Therefore, it is important for organizations to conduct threat modeling on specific cyber-physical systems at the appropriate level. They should perform more detailed threat modeling than what can be supported by using the Purdue model.

The U.S. Food and Drug Administration, in collaboration with MITRE and the Medical Device Innovation Consortium (MDIC), has provided guidance on threat modeling with their Playbook for Threat Modeling Medical Devices. "The playbook is agnostic about specific methodologies, and instead focuses on the

values and principles articulated in the [threat modeling] manifesto and illustrates how different methodologies can be used, alone or in combination, to answer those four key questions." [14]

In How to Develop a Security Vision and Strategy for Cyber-Physical Systems, Gartner provides an approach to developing a security strategy for CPS. One component of that approach is doing a current-state assessment with the aim of getting a clear picture of the situation on the ground, including understanding cyber-physical threats. Other Gartner research in this area includes 3 Initial Steps to Address Unsecure Cyber-Physical Systems and Market Guide for CPS Protection Platforms.

Artificial Intelligence/Machine Learning (AI/ML)

With the emergence and popularity of various artificial intelligence (AI) and machine learning (ML) capabilities, organizations that are using or developing these technologies will need to be proactive in understanding potential AI/ML threats. Some threats to consider include:

- **Data poisoning:** the deliberate insertion of malicious data in the training set. The purpose of data poisoning is to cause an ML model to learn from malicious data, leading to wrong results in the model's output.
- **Prompt injection:** the ability to insert hidden instructions or context into an application's conversational or system prompt interfaces, or in the generated outputs.
- **Membership inference:** the attempt to discover whether a certain data point is part (that is, a member) of the training set. This is a risk when the training data contains sensitive, private or confidential data, such as health information of individuals.
- **Model extraction:** obtaining detailed information about the target ML model. The attacker will then use this information to re-create a close approximation (or "copy") of the target model.
- **Evasion:** an attempt to evade the model's output by querying the model with manipulated input. This is a risk because this leads to unreliable model output.

While the above focus on adversarial threats, organizations must look at other areas where risk could impact the systems they develop or use. These could include:

- **Transparency risks:** There are models that are not explainable. Even vendors don't understand everything about how they work internally.
- **Reliability risks:** The output of a model may contain untruths as well as truths, and these may be retrieved or synthesized from the training data or the context of prompts. Outputs generated by ChatGPT and GenAI, for example, should be validated and moderated to ensure quality assurance is integral to use.
- **Bias risks:** Model training is based on potentially biased data selectively sourced from the web and books. Enterprises must have in place policies or controls to detect biased outputs and deal with them in a manner consistent with their policies and any relevant legal requirements.
- **Intellectual property (IP) and copyright risks:** In the specific case of the ChatGPT application (not the direct API access model), there are currently no verifiable data governance and protection assurances regarding confidential enterprise information. Users should assume that any data or queries they enter into ChatGPT will become public information. Enterprises should take steps to avoid inadvertently exposing enterprise IP. Users must also filter out copyrighted materials in ChatGPT outputs to which OpenAI does not have the rights in order to create potential legal risks. The lack of reliable source references makes this a serious challenge.

Some of the techniques listed in this research can be used to understand the threats and risks that may come from AI/ML systems (whether being developed or used). For example, using sci-fi threat modeling may give an organization a way to speculate about what-if scenarios regarding these systems. With the fast development and increasingly wide use of GenAI and large language models (LLMs), any speculation today is likely to become reality tomorrow. Also, abuser stories could be written about some of the threats listed above.

Cyber shadows loom,

Unseen perils seek their way,

Vigilance shall bloom.

— ChatGPT output responding to the prompt: "Write a threat modeling haiku."

As AI models and techniques become more prevalent, more structured threat modeling approaches will develop. MITRE has a knowledge base of tactics and techniques called the Adversarial Threat Landscape for Artificial-Intelligence Systems (ATLAS), and PLOT4ai is a framework based upon LINDDUN that focuses on privacy-based AI threats.

PLOT4ai

Privacy library of threats for artificial intelligence (PLOT4ai), which is inspired by LINDDUN, provides a library of AI threats along with a threat modeling methodology. The library is divided into the following categories (some of which come from LINDDUN):

- Technique and processes
- Accessibility
- Identifiability and linkability
- Security
- Safety
- Unawareness
- Ethics and human rights
- Noncompliance

The PLOT4ai threat model uses a set of cards, organized by categories, each of which asks a question about a possible threat (for example, “Are we protected against model sabotage?”). Each card has information on how to answer the question and what action may need to be taken. A data flow diagram is represented using four stages where the threat could apply. These stages are design, input, modeling and output.

Strengths

Treating threat modeling as an art enables security and risk management technical professionals to:

- Create a palette of approaches for doing threat modeling.
- Instill a threat-conscious mindset into various roles within the organization.
- Provide approaches to threat modeling in different domains.

- Support secure-by-design and “paved path” principles for development teams.

Weaknesses

While threat modeling has its benefits, there are some cautions that security professionals should consider:

- Heavy cognitive loads require people to think about threats beyond their primary roles.
- Multiple approaches can create confusion as to who is responsible for threat modeling and what artifacts are needed.
- Focusing too much on nonframework approaches could put an organization at risk of compliance violations.

Guidance

Security and risk management technical professionals should:

- Use standard frameworks as a basis for understanding threats in a system. Determine the focus of a threat model (threat view, risk, privacy, asset view, etc.) and use an appropriate framework.
- Involve other roles in the threat modeling process. Using frameworks as a start can educate these roles on the threat modeling process, provide insights into the process and help them better understand the big picture of security. Which roles to include first will depend on the specific threats being explored. Bringing in developers and technical architects will be necessary when conducting threat modeling on an application or production system. Including the legal team may be necessary when modeling threats concerning privacy. If performing threat modeling on a cyber-physical system, then including appropriate device and control system experts will be necessary.
- Start using informal techniques (as mentioned in this research) to instill a threat-conscious mindset in the organization. Work with security champions to promote these techniques, and also learn new techniques from them.
- Wherever there is digital risk within the organization, use some form of threat modeling to understand the extent of the risk.

- Create architectural standards and policies based upon threat modeling results. Standards could include technical controls such as MFA, API security and/or bot mitigation solutions, and service mesh technology for mutual authentication. Other standards could include policies for handling sensitive data like personally identifiable information (PII) or IP.
- Revisit formal threat models on a regular basis to ensure they stay relevant and that implemented controls are still appropriate.

Acronym Key and Glossary Terms

AI	artificial intelligence
API	application programming interface
CAIRIS	Computer Aided Integration of Requirements and Information Security
CCPA	California Consumer Privacy Act
CIA	confidentiality, integrity and availability
CISA	Cybersecurity and Infrastructure Security Agency
CPS	cyber-physical systems
DDoS	distributed denial of service
DFD	data flow diagram
ERP	enterprise resource planning
GDPR	General Data Protection Regulation
GenAI	generative AI
GPS	Global Positioning System
HIPAA	Health Insurance Portability and Accountability Act
HR	human resources
ICS	industrial control systems
IIoT	industrial Internet of Things
IoT	Internet of Things
IP	intellectual property
IR	incident response
LLM	large language model
MDIC	Medical Device Innovation Consortium
MFA	multifactor authentication
ML	machine learning
MOAM	model once, apply to many

OCTAVE	Operationally Critical Threat, Asset, and Vulnerability Evaluation
OT	operational technology
PASTA	Process for Attack Simulation and Threat Analysis
PII	personally identifiable information
PLC	programmable logic controller
ROI	return on investment
SaaS	software as a service
SCADA	supervisory control and data acquisition
SCE	security chaos engineering
SDLC	software development life cycle
WAF	web application firewall

Evidence

[1] Threat Modeling Manifesto

[2] Shifting the Balance of Cybersecurity Risk: Principles and Approaches for Security by Design and Default, Cybersecurity and Infrastructure Security Agency (CISA).

[3] Elevation of Privilege Game, Shostack + Associates.

[4] K. Shortridge, Creating Security Decision Trees With Graphviz.

[5] Dark Hammer: A Retrospective of Science Fiction Prototyping, Threatcasting Lab, Arizona State University.

[6] APA Dictionary of Psychology, American Psychological Association.

[7] K. Shortridge and A. Rinehart, Security Chaos Engineering: Sustaining Resilience in Software and Systems, O'Reilly Media, 2023.

[8] STELLA: Report from the SNAFUcatchers Workshop on Coping With Complexity, SNAFUcatchers

Consortium for Resilient Internet-Facing Business IT.

[9] Application Security Podcast: Interview With Wolfgang Goerlich, Security Journey.

[10] Bots Ruin Everything: How Taylor Swift Concert Made Us Compete Against Nonhumans, Cybernews.com.

[11] As Online Gambling Grows, So Does the Financial Crime Risk, The Wall Street Journal.

[12] Five Ways the Gaming & Gambling Industry Is Targeted by Bad Bots, Imperva.

[13] M. Krotofil, Industrial Control Systems: Engineering Foundations and Cyber-Physical Attack Lifecycle, May 2023.

[14] Playbook for Threat Modeling Medical Devices (PDF), MITRE.

Recommended by the Author

Some documents may not be available as part of your current Gartner subscription.

Market Guide for Security Threat Intelligence Products and Services

Use Adversary-Generated Threat Intelligence to Improve Threat Detection and Response

5 Frequently Asked Questions About Threat Modeling

Implement a Continuous Threat Exposure Management (CTEM) Program

3 Ways to Apply a Risk-Based Approach to Threat Detection, Investigation and Response

Top 5 Priorities for Managing AI Risk Within Gartner's MOST Framework

4 Ways Generative AI Will Impact CISOs and Their Teams

How to Securely Design and Operate Machine Learning

Table 1: Threat modeling frameworks

Framework	Description	Focus
Attack trees/decision trees	Model attacks from the perspective of an attacker.	Attacker
DREAD	Risk assessment based upon damage, reproducibility, exploitability, affected users and discoverability.	Risk impact
LINDDUN/PLOT4ai	Catalogs of privacy threats for information and artificial intelligence/machine learning (AI/ML) systems.	Privacy
Operationally Critical Threat, Asset, and Vulnerability Evaluation (OCTAVE)	Used to identify and assess risks to IT assets. It assesses organizational risks, not technological risks.	Asset risk
Process for Attack Simulation and Threat Analysis (PASTA)	A seven-step, risk-centric modeling framework that focuses on the business impact of threats.	Business risk
STRIDE	Identifies threats based upon spoofing, tampering, repudiation, information disclosure, denial of service and elevation of privilege.	Threats
Trike	Used for security auditing processes to identify whether an assigned level of risk is “acceptable” for an asset.	Audit risk

Framework	Description	Focus
Visual, Agile, Simple Threat (VAST) modeling	Covers the entire software development life cycle (SDLC) across an organization using the three pillars of scalability: automation, integration with SDLC tools and collaboration with cross-functional stakeholders.	SDLC risk

Source: Gartner (November 2023)

Table 2: Examples of User and Abuser Stories

User Story	*Abuser Story*	*Socratic Question*
As a payment app user, I want to transfer money from my account to the account of another app user.	As a malicious user, I want to transfer money from a legitimate user's account to my account without their consent.	What do you really know about your users? Do all of them have good intentions? Is authentication enough?
As a commenter, I can add comments to a blog post.	As a malicious commenter, I can add JavaScript to a blog post comment that executes a malicious function in a blog reader's browser.	Are you sure you want just anyone to add comments? Should we be maintaining the comments section or would it be better to use a third-party service?
As an employee, I can enter expense information into the reporting system to get reimbursed for travel expenses.	As a malicious employee, I can falsify expense records to get money transferred to my account.	What do you really know about our employees? Do all of them have good intentions?

Source: Gartner (November 2023)