K2's AI - Security And Privacy Issues

Learning Objectives

Differentiate between AI security and privacy risks

List examples of key security and privacy risks associated with many of today's AI platforms

Identify best practices for mitigating the security and privacy risks associated with AI

Major Topics

Major provisions of security and privacy policies

Key risks associated with using AI tools

Steps to take to mitigate AI security and privacy risks

Safest AI Tools – Constitution/Indemnification

Safety

- Claude
- Copilot
- Perplexity AI
- ChatGPT
- Gemini
- Llama (Meta)

Popularity

- ChatGPT - 61% (14% in India)
- Copilot – 14%
- Gemini – 13%
- Perplexity AI – 6.2%
- Claude – 3.2% (80% business)
- All others below 1%

Are There Dangers With AI? ***Absolutely***!

Biased Outputs	Copyright infringement	Data privacy
Data security	Deepfakes	Hallucinations

Are There Dangers With AI? ***Absolutely***!

Biased Outputs	Copyright infringement	Data privacy
Data security	Deepfakes	Hallucinations

Copyright Infringement

Who owns the copyright to materials generated by AI?

Response, Per ChatGPT	Response, Per Gemini
Generally, the entity or individual that creates the AI-generated content is considered the copyright owner. However, this may vary based on jurisdiction and specific contractual agreements	*In the United States, the Copyright Office has stated that it will not register works that were created by an autonomous AI tool. This means that, under current US law, AI-generated works are either in the public domain or they are derivative works of the materials that the AI was trained on*

Copyright Infringement

New York Times Has Sued MSFT & OpenAI!

- In December, NYT sued Microsoft and ChatGPT's parent company, OpenAI
- The Times alleges "*mass copyright infringement*"
- The Times continues with *"These tools were built with and continue to use independent journalism and content that is only available because we and our peers reported, edited, and fact-checked it at high cost and with considerable expertise"*
- One example provided by the Times OpenAI's software producing almost identical text to a Times article about predatory lending practices in New York City's taxi industry

Data Privacy

- Consider the type of data you might upload into an AI tool or platform…how comfortable are you that sensitive data remains private and not identifiable back to an individual, a client, or an organization?
- Other issues include
  - Is there informed consent about collection and use
  - The threat of exposing sensitive data through a breach or attack
  - Lack of transparency regarding using an individual's or organization's data

European Union Privacy Regulations

1. General Data Protection Regulation (GDPR)

- **Still the cornerstone** of EU data privacy law since its enforcement in 2018
- Focuses on:
  - Transparency
  - Accountability
  - Individual rights
- Influenced privacy laws globally, known as the "Brussels Effect."

2. AI Act (Effective August 2024)

- Regulates AI systems based on **risk levels** to fundamental rights, health, and safety
- **Prohibits** high-risk applications like:
  - Social scoring
  - Real-time biometric identification in public spaces
- Requires for high-risk AI:
  - Risk management systems
  - Transparency measures
  - Data governance practices
  - Human oversight
- Full compliance expected by **August 2026**
- Includes upcoming initiatives:
  - AI literacy programs
  - Codes of Practice for General Purpose AI (due May 2025)
  - Governance for systemic-risk AI models[1]

3. Digital Operational Resilience Act (DORA)

- Complements GDPR and AI Act by focusing on **resilience in digital operations**, especially in financial services.
- Ensures robust cybersecurity and incident reporting

4. GDPR Simplification Proposal (May 2025)

- Targets **small mid-cap enterprises (SMCs)**:
  - Defined as having <750 employees, balance sheet <€129M, turnover <€150M
- Extends derogations from record-keeping obligations (Article 30 GDPR)
- Aims to **reduce administrative burden** while maintaining data protection standards
- Encourages tailored codes of conduct and certification mechanisms for SMCs[2]

5. New Procedural Rules for GDPR Enforcement

- Adopted in May 2025 to improve enforcement in **large, cross-border cases**
- Designed to make GDPR enforcement **faster and more effective**[3]

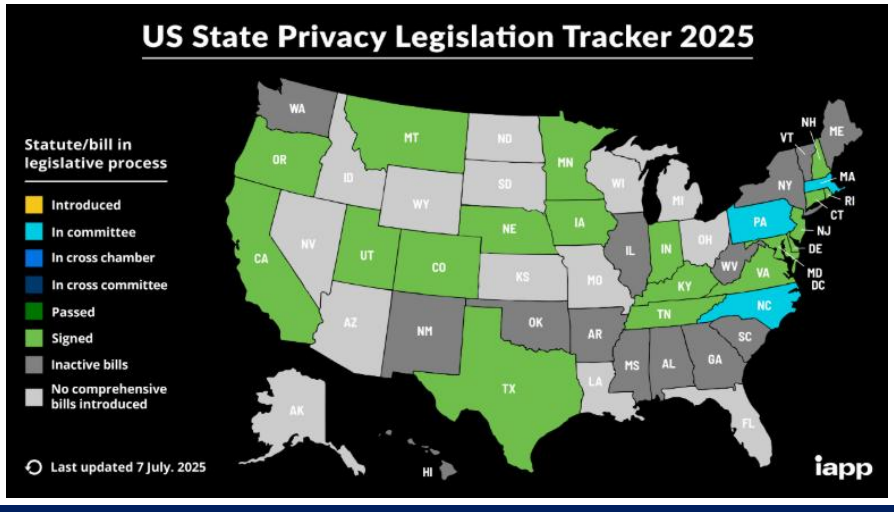
Canadian Privacy Regulations

- Federal Privacy Laws
  - Privacy Act
    - Governs how federal government institutions collect, use, and disclose personal information
    - Applies to federal departments, agencies, and Crown corporations
  - PIPEDA (Personal Information Protection and Electronic Documents Act)
    - Applies to private-sector organizations engaged in commercial activities across Canada.
    - Covers the collection, use, and disclosure of personal information during commercial activities
    - Enforced by the Office of the Privacy Commissioner of Canada[1]
  - Bill C-27 (Proposed)
    - Introduces the Consumer Privacy Protection Act (CPPA) and the Artificial Intelligence and Data Act (AIDA)
    - Aims to modernize PIPEDA and regulate AI systems
    - Still under legislative review as of 2025[2]
- Health Sector Privacy Laws
  - Personal Health Information Protection Act (PHIPA) – Ontario
  - Health Information Act (HIA) – Alberta
  - An Act Respecting Health Services and Social Services (ARHSSS) – Quebec
  - These laws govern the handling of personal health information by healthcare providers and institutions
- Provincial and Territorial Privacy Laws. Some provinces have their own private-sector privacy laws deemed substantially like PIPEDA:
  - Quebec: Act Respecting the Protection of Personal Information in the Private Sector (amended by Law 25 in 2023)
  - British Columbia: Personal Information Protection Act (BC PIPA)
  - Alberta: Personal Information Protection Act (AB PIPA)
- Employment and Sector-Specific Laws
  - Employment-related privacy protections may be embedded in labor codes or sector-specific regulations
  - Some provinces have additional rules for public bodies and educational institutions

State Privacy Regulations (16 Now)

1. California Consumer Privacy Act (Oct 1, 2020, amended Jan 1, 2023)
2. Colorado Privacy Act (Jul 1, 2023)
3. Connecticut Personal Data Privacy and Online Monitoring Act (Jul 1, 2023)
4. Delaware Personal Data Privacy Act (Jan 1, 2025)
5. Indiana Consumer Data Protection Act (Jan 1, 2026)
6. Iowa Consumer Data Protection Act (Jan 1, 2025)
7. Kentucky Consumer Data Protection Act (Jan 1, 2026)
8. Maryland Online Data Privacy Act (Oct 1, 2025)
9. Minnesota Consumer Data Privacy Act (Jul 31, 2025)
10. Montana Consumer Data Privacy Act (Oct 1, 2024)
11. Nebraska Data Privacy Act (Jan 1, 2025)
12. New Hampshire SB 255 (Jan 1, 2025)
13. New Jersey SB 332 (Jan 15, 2025)
14. Oregon Consumer Privacy Act (Jul 1, 2024)
15. Rhode Island Data Transparency and Privacy Protection Act (Jan 1, 2026)
16. Tennessee Information Protection Act (Jul 1, 2025)
17. Texas Data Privacy and Security Act (Jul 1, 2024)
18. Utah Consumer Privacy Act (Dec 31, 2023)
19. Virginia Consumer Data Protection Act (Jan 1, 2023)

US State Privacy Legislation 2025



UT Digital Choice Act HB418 7/1/26

- Utah’s groundbreaking new Digital Choice Act will finally give people agency over their data on social media
- Under the act, individuals will be able to use open-source protocols to seamlessly move their content and relationships to new apps if they are unhappy with the experience on a social media site. It aims to challenge Big Tech's monopolistic practices and foster competition in the digital ecosystem. The law reflects growing public sentiment for greater control over personal data and targeted advertising, marking a significant step toward a more user-centric internet. Provisions:
  1. Complete Data Portability
  2. Right to Deletion
  3. Enforces Interoperability

CA SB 53 Transparency In Frontier Artificial Intelligence Act (TFAIA)

- California Governor Gavin Newsom signed SB 53 on 9/30/25, the nation’s first law placing guardrails on frontier AI
- Newsom described it as installing “commonsense guardrails” to build public trust without stifling innovation. Anthropic, Stanford, UC Berkley, and others endorsed
- The new law applies to frontier AI models, or the most advanced systems with significant capabilities and risks. What SB53 does:
  1. **Transparency:** Requires major AI developers to publish a framework showing how they follow national and international standards, making their practices visible to the public
  2. **Innovation:** Establishes CalCompute, a state-backed public computing consortium to support safe, equitable AI research
  3. **Safety:** Sets up a system for companies and the public to report critical AI safety incidents directly to California’s Office of Emergency Services
  4. **Accountability:** Protects whistleblowers who expose serious risks from frontier models and imposes civil penalties for noncompliance, enforceable by the attorney general
  5. **Responsiveness:** Directs the Department of Technology to recommend annual updates, so the law keeps pace with rapid advances and international standards

Data Security

- Like other large caches of data, the data stored in AI tools and platforms can present an inviting target because of the private nature of the data
- Additionally, a general lack of human oversight can increase the vulnerability of the data
- Moreover, the complexity of AI systems presents challenges concerning storing and protecting sensitive data
- In fact, as reported on 11/10/23, Microsoft temporarily blocked team members from using ChatGPT due to security concerns

Understand what you're getting into with generative AI

SOFTWARE LICENSES AND PRIVACY POLICIES

Just Another Battle In A Long War...

- Historically, software licensing and use policies have been a problem for parties on both sides of the fence
- That hasn't changed with AI, and likely will not change soon
- On the developer side, there is an incentive to get as much quality data as possible into the LLMs, while simultaneously building up the subscriber base
- However, on the end-user side, we should not share data that is private/sensitive with the outside world
  - Do you ***really*** want to upload your tax return to an AI platform?

Add AI And It Gets Complicated

- Software licensing and data usage policies have always been hard to understand – even with search engines
- Now, consider the impact of Artificial Intelligence (AI) when added to the mix
- Could creating a prompt asking for tax advice expose your personal data or that of a client?
- Could uploading financial data for a privately-held company to an AI platform cause that data to leak?
- Can we trust what the AI companies are telling us about data security and privacy?

Understanding Licenses Is Critical

- To better understand the risks associated with privacy and AI, we need to **READ** and **UNDERSTAND** the major documents associated with the application
- Two of the key documents associated with your rights/license to use software or cloud-based services include:
  - End User License Agreement (EULA) a/k/a Terms of Service (ToS)
  - Privacy Policy

ChatGPT's Terms Of Use Key Points

- Must be at least 13 and, if under 18, have parental consent
- Cannot use ChatGPT for “illegal, harmful, or abusive” activity, including infringing upon or violating anyone’s rights
- You cannot represent that results are human-generated
- You retain ownership rights to your inputs and outputs
- ChatGPT can use your content to “provide, maintain, develop, and improve our services”
- You can opt out of ChatGPT using your data to train ChatGPT

https://openai.com/policies/terms-of-use/

What Data Is Used To Train LLMs

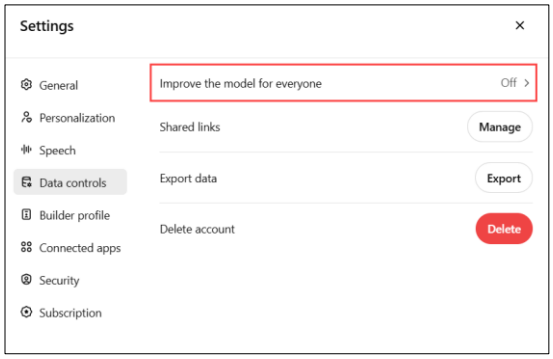
Information that is publicly available on the internet

Information that we partner with third parties to access and use

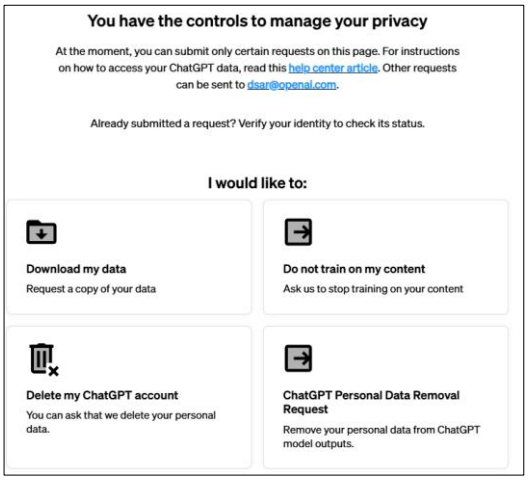
Information that our users or human trainers and researchers provide or generate

To Opt-Out Of ChatGPT's Training…

- Visit ChatGPT's **Settings**
- Next, click **Data controls**
- Finally, turn off the **Improve the model for everyone** option as shown on the right
- It's a good idea to review all the other settings also to ensure they align with your expectations and needs



ChatGPT's Privacy Request Portal



- You can find the **Privacy Request Portal** at https://k2e.fyi/OAIPrivacy or by using the QR code below

Privacy Policy

- Document which discusses a company's privacy policy, including data like
  - What data is gathered from you
  - How your information will be used
  - The purpose of gathering that data and how your data may be used
  - When your data will be disclosed to others, including sensitive personal data and information
    - Not always specific on with whom it will be shared
  - Security policies and procedures
- Consider ChatGPT's Privacy Policy

Acceptable Use Policy

- Just as the EULA/ToS, and privacy policy spell out the rights and responsibilities of both parties in a relationship between an end user and a software company, an Acceptable Use Policy (AUP) governs the acceptable use of the organization's technology hardware, software, and data by its employees and contractors
- Although acceptable use policies are not required by law, they help both employees and companies properly set expectations about how these tools will be used (and how they will NOT be used) so all parties know their rights and responsibilities

Acceptable Use Policy Defines

Ownership and general use of technology

Security and handling confidential information

Types of unacceptable use of systems

E-mail and communication guidelines

Blogging and social media

Consequences for violation of these policies

Potential Harms From AI Systems

Source: "Artificial Intelligence Risk Management Framework" (AI 100.1) by US National Institutes for Standards and Technology (NIST)

SO, WHAT ARE THE AI COMPANIES SAYING?

ChatGPT Data Concerns

- OpenAI (ChatGPT's parent company) collects:
  - IP address, location, browser type, date/time, length of session, device name, and operating system
  - Uses cookies to track browsing activities both in chat window and site
  - Records complete transcripts of your prompts and conversations
  - ChatGPT reserves the right record any data you upload, such as an Excel file or a set of financial statements
- As discussed previously, to enhance security and privacy, consider opting out of model improvement and deleting your chat history

Gemini Concerns

- Google publicly stated that Gemini collects the following information about your use of Gemini
  - Conversations, locations, feedback, and usage info
- Regarding accessing to your Gemini conversations, Google says the following: *We take your privacy seriously, and we do not sell your personal information to anyone. To help Gemini improve while protecting your privacy, we select a subset of conversations and use automated tools to help remove user-identifying information (such as email addresses and phone numbers).*

Gemini Concerns

- Google goes on to state: *Please don't enter confidential information in your conversations or any data you wouldn't want a reviewer to see or Google to use to improve our products, services, and machine-learning technologies.*
- *Gemini Apps conversations that have been reviewed by human reviewers are not deleted when you delete your Gemini Apps activity because they are kept separately and are not connected to your Google Account. Instead, they are retained for up to three years.*

MICROSOFT'S COPILOT OFFERING IS A BIT DIFFERENT

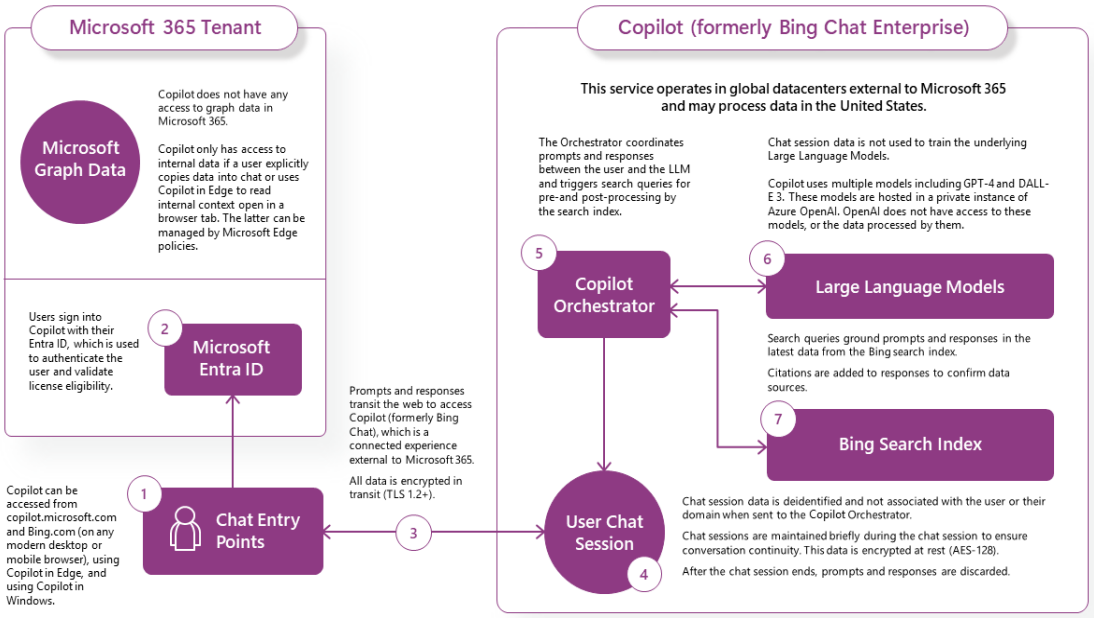
Microsoft's AI Copilots

Name	Products	Monthly Cost	Commercial Data Protection Included?
Copilot for MS 365 (Business/Enterprise)	Microsoft 365 apps (Word, Excel, PowerPoint, Outlook, Teams)	$30 per user	Yes
Copilot in Windows (Bing Chat)	Windows OS	Free	Not for home users, included with most business/enterprise O365/M365 plans
Copilot Pro for Individuals	Advanced features on top of standard Copilot, plus integration with home Microsoft 365 apps	$20/user/mo.	Not specified
Copilot for Security	Microsoft's cybersecurity products	Consumption-based fee - $4/hour	Not specified
Copilot for Finance, Sales, and Service	Financial operations, sales optimizations, service enhancements	$50/user/mo., $20/user/mo. if already have MS 365	Not available to other customers, runs on Microsoft cloud in separate instance of ChatGPT, not used by MS to train models by default
Designer for Copilot	Image creation and editing	Not available	No
Copilot GPTs and Azure AI Studio	Custom generative AI assistants and solutions	Not specified	Not specified

Commercial Data Protection

- **Authentication**: It uses Microsoft Entra ID to ensure only authorized users can access Copilot with commercial data protection
- **Data Encryption**: Chat data is encrypted both in transit and at rest using advanced encryption standards
- **Privacy**: Prompts and responses are not saved or used to train the underlying models//Microsoft has no "eyes-on" access to this data
- **Anonymity**: Search queries triggered by prompts are not linked to users or organizations
- **Ad Protection**: Advertising shown to users isn't targeted based on their work identity or chat history

Copilots and Data Protection

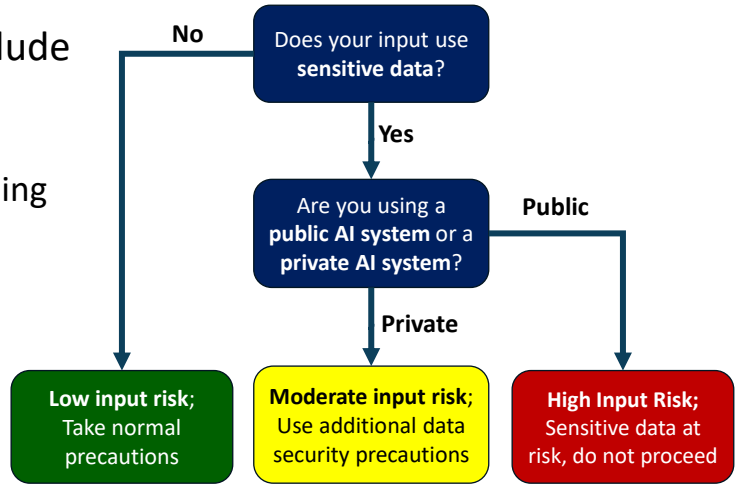


Three Types Of GenAI Risk

Input risks	Output risks	System risks
• The risks associated with exposing your proprietary data to an AI system and that data becoming compromised, which results in unauthorized disclosure of confidential information	• The risks that the outputs from the AI model will be low quality, inaccurate, or incomplete and the models lose their integrity based on including an unacceptable number of erroneous data points	• Risks associated with the servers hosting the AI system being compromised and the data model needs to be recovered from backups

Gen AI Input Risk Decision Tree

- Types of input attacks include attempts to:
  - Crash AI model
  - Exfiltrate memorized training data

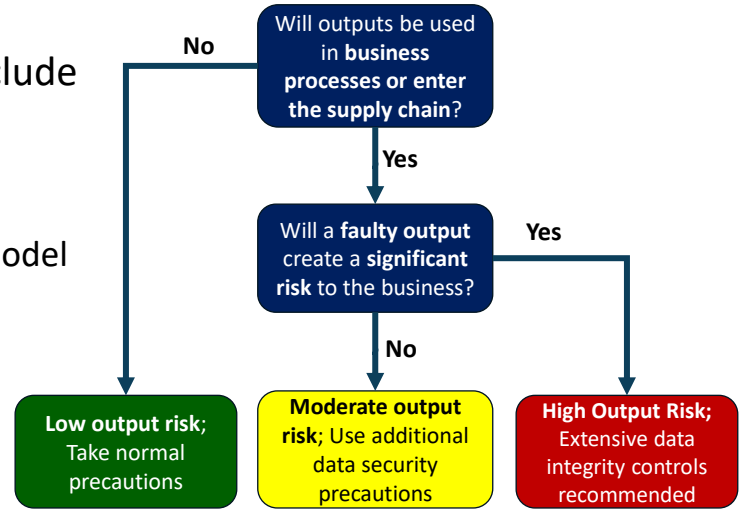


Source: "Address Security & Privacy Risks for Generative AI" by InfoTech Research Group

Gen AI Output Risk Decision Tree

Types of output attacks include attempts to

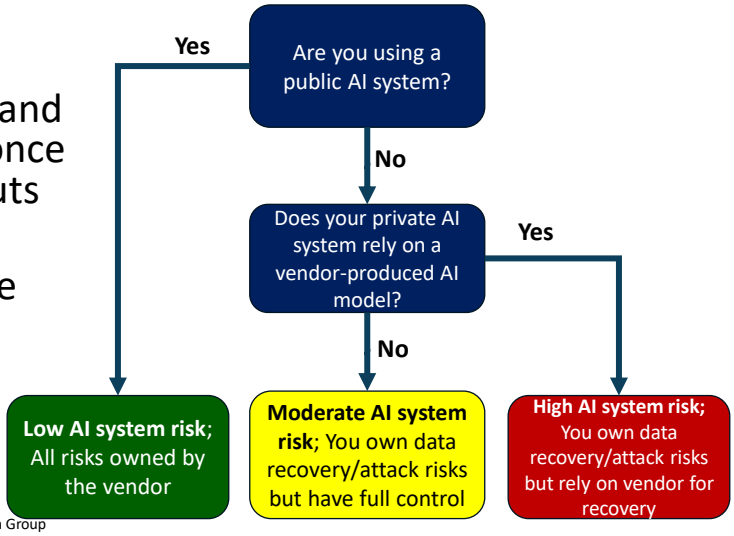
- Data poisoning that could corrupt AI output
- Weaponization of an AI model
- Sponging to slow down processing speed of AI model



Source: "Address Security & Privacy Risks for Generative AI" by InfoTech Research Group

Gen AI System Risk Decision Tree

- Public AI systems like ChatGPT, Google Gemini, and Bing chat are developed once and trained based on inputs from many users
- Private AI systems must be trained, managed, and maintained by employees using internal organization resources



Source: "Address Security & Privacy Risks for Generative AI" by InfoTech Research Group

Summary And Wrap Up

- Proactively manage the security and privacy settings in the AI platforms that you use
- Never enter or upload private or sensitive data into your generative AI prompts
  - Scrub names, SSN's, account numbers, etc. from your data before uploading it to a generative AI tool
- Train your team members on how to use these tools safely and securely…remember the survival of your business may depend upon it!