SECURANCE CONSULTING
the advantage of insight

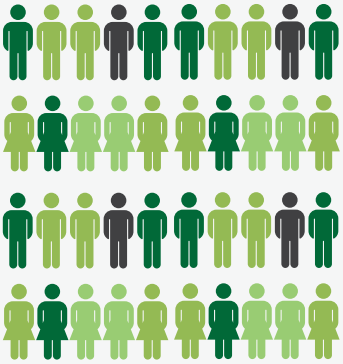
Proprietary ChatGPT: Enterprise AI Security Risks and Solutions for a Smarter World

www.securanceconsulting.com

INTRODUCTION

Artificial intelligence (AI) is a transformative field that has rapidly emerged as a driving force behind numerous technological advancements in recent years, such as natural language processing, cybersecurity, machine learning, and robotics. Most notably, OpenAI's chatbot, ChatGPT, has become increasingly popular among enterprises due to its effectiveness, accuracy, and efficiency. Powered by advanced natural language processing and machine learning techniques, proprietary ChatGPT technologies bring value to enterprises by streamlining communication, creating images, writing code, composing documents, conducting research, and more. With its ability to analyze vast amounts of data, make predictions, and adapt to new information, ChatGPT's power to shape the future is boundless, offering opportunities and challenges that have the potential to redefine the way we live, work, and interact with technology.

2023 OpenAI and ChatGPT Statistics[1]



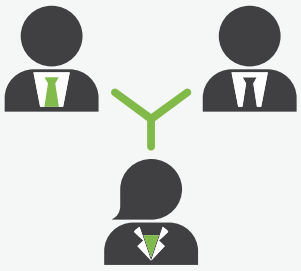
ChatGPT has more than **100 MILLION** users, as of June 2023

53%

53 percent of readers believe that ChatGPT-generated content has been created or edited by humans, specifically in the fields of finance, health, technology, travel, and entertainment

OpenAI software is used by 902 companies across a variety of business sectors. 27 percent of these companies are technology-focused

Over 3 percent of employees have entered confidential company data into ChatGPT



Over 8 percent of company employees have tried using ChatGPT in the workplace[2]

By 2025, 30 percent of outbound marketing messages from enterprises will be synthetically generated[3]

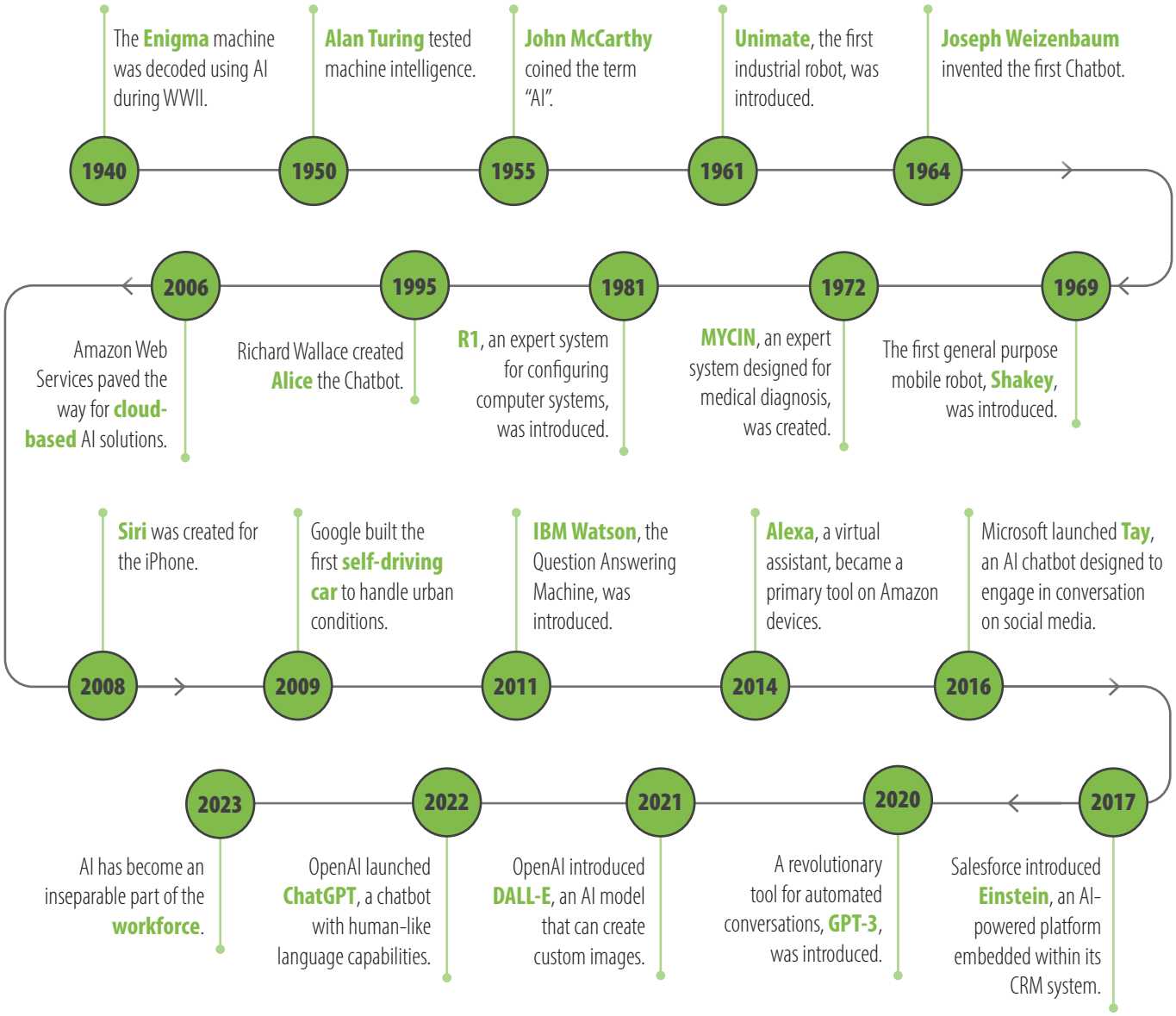
> ChatGPT has emerged as a powerful tool for various use cases. Enterprises can customize and fine-tune large language models like ChatGPT, using their own data to create more efficient and domain-specific business tools.[4]

WHAT IS CHATGPT?

To fully understand what ChatGPT is and how it works, we must first look at the concept of AI as a whole. In its simplest form, AI is a field that combines computer science and robust datasets to enable problem-solving. It also encompasses machine learning and deep learning, which consist of AI algorithms seeking to create expert systems that make predictions or classifications based on input data. AI also refers to large language models (LLMs) that can take raw data and "learn" to generate statistically probable outputs when prompted.[5]

Over the years, there have been many AI models. ChatGPT is one of the latest LLMs that has facilitated significant improvements in the performance of AI and its ability to automate certain processes, secure critical assets, and integrate into businesses' workflows.

Timeline of Artificial Intelligence

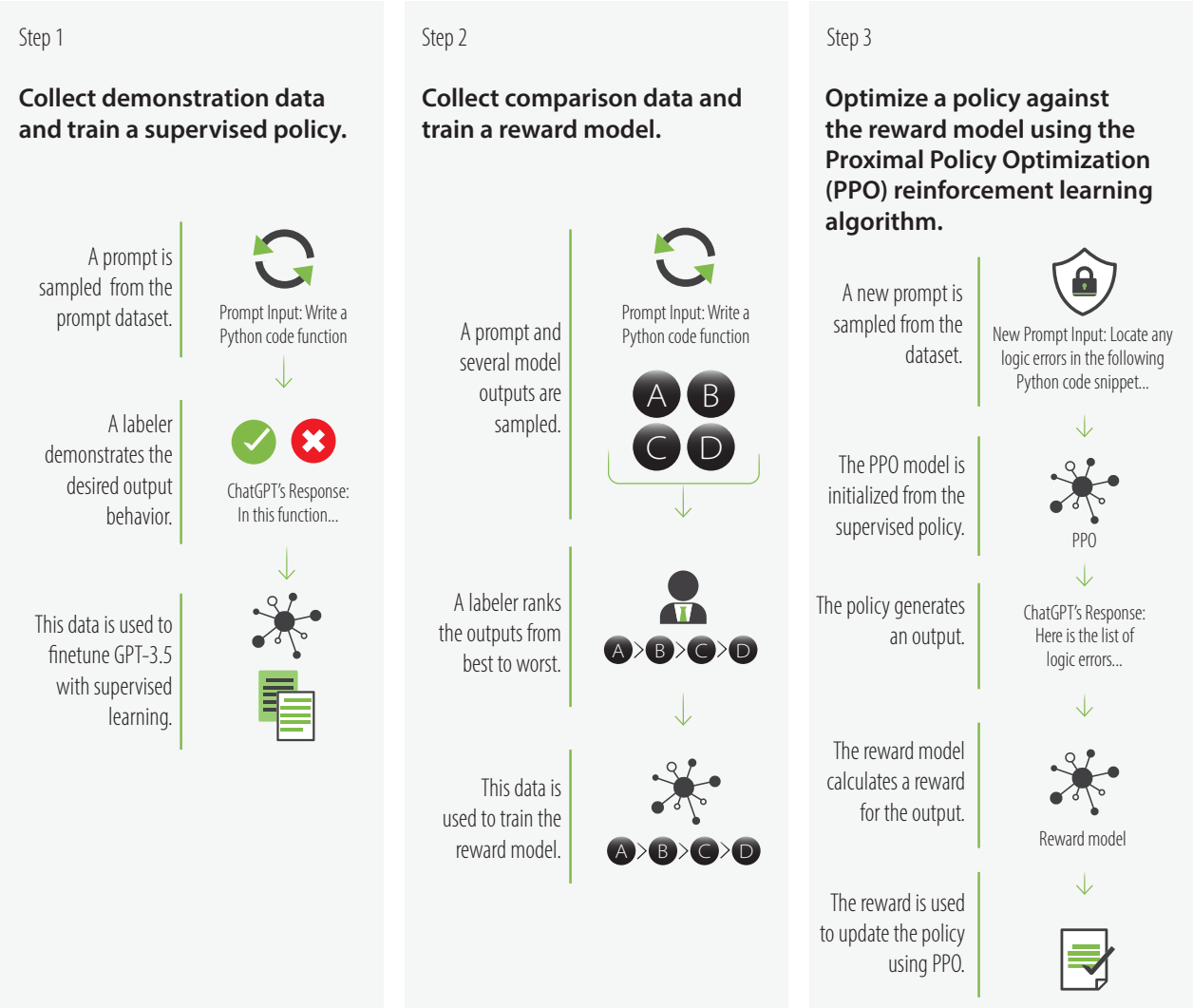


Now that we have covered what the term AI means, we can take a deeper dive into ChatGPT. ChatGPT is an AI-driven chatbot created by the research company OpenAI. ChatGPT works on the basis of dialogue to understand the user's intent, grasp the context, and answer various questions by drawing information from a large pool of data. ChatGPT is not the only chatbot, but it is one of the most accurate and accessible chatbot tools.

What makes ChatGPT different from other forms of AI? ChatGPT uses a new LLM architecture called the Generative Pre-Trained Transformer (GPT). The chatbot is trained on a large set of text data to produce text, recognize patterns, and understand the inner structure of language.[6] ChatGPT is also considered revolutionary in that the conversation between users and the AI feels more natural.

OpenAI's GPT series consists of the original GPT model, and the following iterations: GPT-2, GPT-3, ChatGPT, and GPT-4, with GPT-4 being the latest AI system in the series, as of March 2023. Each model follows the research path from the previous version based on OpenAI's deep learning approach. Many enterprises are using the underlying LLM of GPT-3 (the baseline for the ChatGPT variant) and other generative AI models as foundational tools to develop proprietary AI technologies for various tasks, such as content generation, customer support, personalized product recommendations, and training | education platforms. By using GPT-3 as a starting point, businesses can create unique AI tools that align with their goals, industry-specific needs, and target audience, providing them with a competitive edge and the ability to deliver innovative and valuable services.

How Does ChatGPT Work?[7]



CHATGPT USE CASES AND APPLICATIONS

Below are some examples of how enterprises are using proprietary ChatGPT technologies.

PrivateGPT

PrivateGPT is a privacy layer for LLMs, such as ChatGPT. Unlike Public GPT, which caters to a wider audience, PrivateGPT is tailored to meet the specific needs of individual organizations, ensuring the utmost privacy and customization. With PrivateGPT, organizations can redact sensitive and confidential data, such as social security numbers, credit card information and other personally identifiable information (PII). Plus, PrivateGPT does not store user information on its servers and does not track usage, ensuring that the data remains confidential and secure.

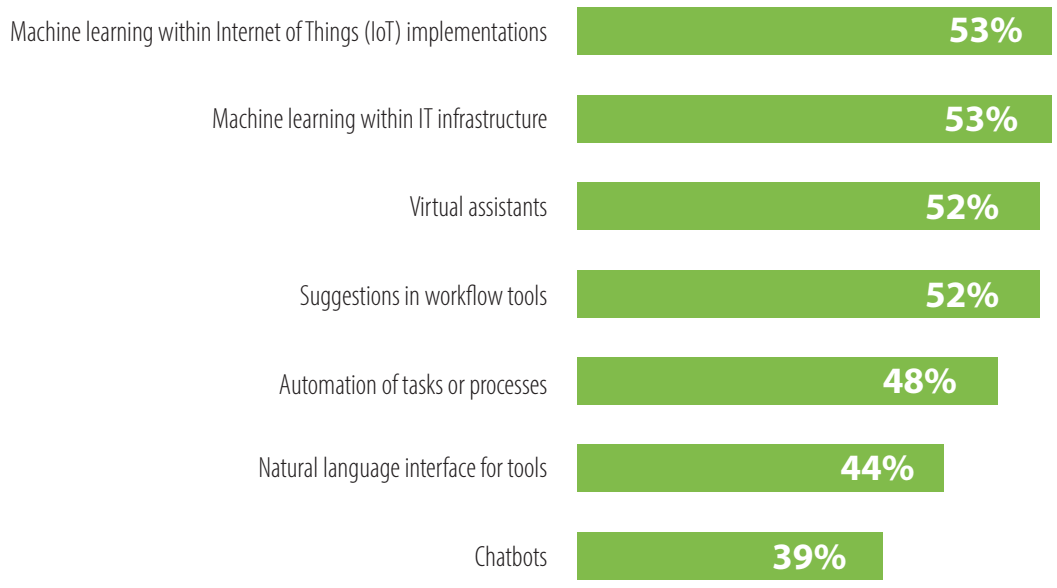
PrivateGPT empowers enterprises to create personalized GPT-3 models without the need for coding or technical expertise. This allows organizations to tailor the models to specific requirements, enhancing their effectiveness and relevance.[8] It is important to note that PrivateGPT models require dedicated resources and expertise to set up, maintain, and continuously update. However, for enterprises with stringent data privacy requirements or specific customization needs, PrivateGPT models offer a solution that combines the power of AI with enhanced control and security.

Data Security

Developing new tools and tactics to defend against cybercrime is a continuous battle, and, with a massive shortage of skilled cybersecurity professionals, IT teams are struggling to keep up with the volume of threats. To ease this burden, organizations are turning toward applications, such as ChatGPT, to simplify the labor-intensive activities of sorting through and identifying malicious activity, giving back valuable time to the IT staff.

Security professionals can now use proprietary ChatGPT technologies to filter the data collected by security tools for malicious activity by entering queries in plain text, without the tool needing to understand the underlying database structure. This approach eliminates the need to use programming languages, such as SQL, and helps reduce time spent searching for threat data manually.[9]

Common Ways Enterprises Use AI[10]



Proprietary Training Platforms

Businesses can leverage ChatGPT to develop AI-powered training and education platforms. These platforms can offer interactive and personalized learning experiences, simulate real-world scenarios, and provide instant feedback, enhancing employee training, on-boarding processes, and customer education programs.

Analyzing Legal and Regulatory Requirements

By processing and analyzing large amounts of data, the enterprise's AI system can analyze lengthy legal documentation, and provide insights, predictions, and recommendations, streamlining the review process and saving the company valuable time. Additionally, proprietary AI models can automate aspects of the contract drafting process, such as generating standard clauses and performing basic checks for consistency and formatting. However, it is critical to involve legal professionals to review and finalize contracts to ensure compliance and protect the parties involved.

Marketing

Proprietary GPT models can be utilized to generate high-quality content for marketing efforts, such as articles, blog posts, product descriptions, and images. The model can be trained on the enterprise's existing content to maintain consistency and brand voice while automating the content creation process. Proprietary AI systems can also go a step further by analyzing customer preferences, user behavior, analytics, and purchase history to deliver personalized product recommendations. These systems can help improve selling opportunities, increase customer engagement, and enhance the overall shopping experience.

Fraud Detection and Prevention

Proprietary AI models can be used as part of a comprehensive fraud prevention strategy. With its ability to process and analyze data in real-time, the model can raise alerts or trigger automated responses when it identifies suspicious patterns or activities that align with known fraud indicators. The most common uses of ChatGPT in assisting with fraud detection include anomaly detection, identifying social engineering scams, real-time monitoring, and pattern recognition.

Public Information

ChatGPT's role is to complement existing communication efforts and provide accessible, prompt, and reliable information to the public. ChatGPT can retrieve relevant information from a wide range of sources, including official websites, government databases, and public records. Users can ask questions about specific topics, such as government services, regulations, public health guidelines, or local events, and ChatGPT can provide accurate and up-to-date information, taking away the need for users to visit multiple websites.

Corporate Email Addresses

ChatGPT uses machine learning algorithms to analyze text, learn how words and phrases relate, and understand human language. ChatGPT's human-like conversations have enabled organizations to use the application for customer support by training it to review and quickly reply to emails. It interacts in a conversational way and can answer follow-up questions, challenge incorrect premises, and even decline inappropriate requests, saving the enterprise time, money, and resources.[11]

RISKS OF ENTERPRISE CHATGPT

With all the uses and benefits of ChatGPT, it almost seems too good to be true, but there are legal implications and misuse cases that make some organizations wary of AI. Your organization should consider all aspects of ChatGPT before implementing any new tool or solution.

ChatGPT Mobile Applications

Enterprises already using proprietary ChatGPT on their desktop may choose mobile applications for faster smartphone access to their system. Although convenient, mobile ChatGPT applications are a concern for enterprises due to the lack of oversight and control over data restrictions. Securing proprietary mobile applications becomes a unique challenge as these new technologies have not yet matured alongside enterprise security configurations. With an online version of ChatGPT, firewalls can monitor the general traffic, and, based on the firewall's configurations, block specific requests if an employee tries to access inappropriate or illegal content. The problem with private mobile applications is that the firewall cannot see the conversation stream, and, therefore, cannot restrict certain information and illicit content. Moreover, employees could mistakenly give away trade secrets or confidential data while using it in their work to generate email responses.

Enterprise Data Leakage

Enterprise use of ChatGPT may result in leakage of sensitive data, customer information, intellectual property, and trade secrets. Because ChatGPT can store chat history and use the conversations to train its models further, if another user requests similar information, ChatGPT may retrieve the proprietary data it stored to answer similar queries from other users. An incident like this happened to Samsung Electronics earlier this year when one of their engineers uploaded internal source code to ChatGPT. Although the severity of the leak remains unclear, Samsung banned the use of generative AI tools across the organization. Other large enterprises, such as Amazon and JPMorgan Chase, have followed suit in restricting or banning the use of ChatGPT. The worry now is that the data companies share with AI chatbots will be stored on servers owned by organizations such as OpenAI, Microsoft, and Google—with no easy way to access and delete it.[12]

> Before a tool becomes available to the public, developers need to ask themselves if its capabilities are ethical. With reputations and revenue on the line, industries must come together to have the right protections in place and make the ChatGPT revolution something to welcome, not fear.[13]

Privacy Concerns

With any new tool or solution that deals with confidential data, it is critical for businesses to be transparent about the types of data stored and how they are used. If customers are unaware that ChatGPT powers an enterprise chatbot and stores their data, legal battles could arise. Even if organizations use a proprietary GPT model, the application houses an enormous amount of sensitive information that could result in a devastating impact to the enterprise and its customers if compromised.

Laws and Regulations

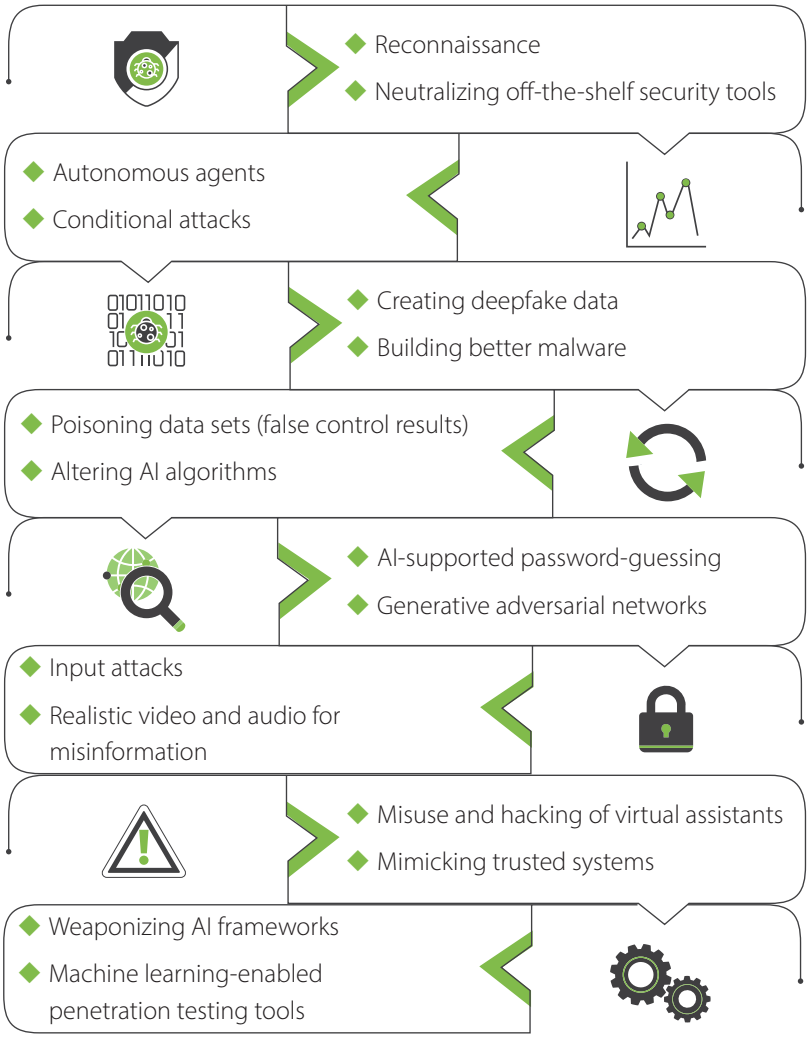
The legal side of AI is rapidly evolving to keep up with the technology landscape. The European Union's (EU's) General Data Protection Regulation (GDPR) establishes "The Right to Explanation," which stipulates that customers impacted by AI have a right to request information about how the AI works. Many industries with strict regulations, such as healthcare and finance, have already developed rules limiting the types of AI that can be used.[14] Violations and noncompliance with certain AI laws and regulations can have devastating consequences for enterprises and their staff, including jail time, ruined reputations, data loss, and hefty fines.

Intellectual Property Disputes

One of the most common questions about ChatGPT is whether the content it creates is owned by the user inputting the commands. Recently, ChatGPT has been used by businesses to create digital marketing content, such as blog posts, articles, and images. It has also been used by programmers to write code. This raises the question of whether the intellectual property belongs to the user or ChatGPT, which could lead to copyright issues, especially if more than one party profits from ChatGPT's assistance.

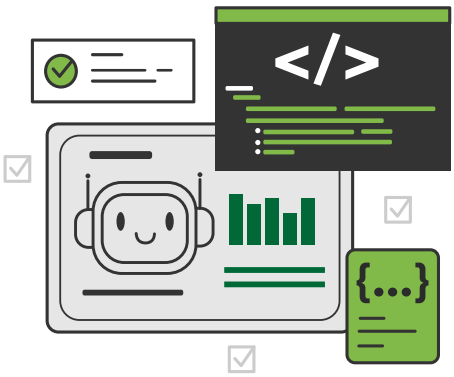
AI Weaponization Tactics

With so many public AI models available, bad actors have found numerous ways to use proprietary ChatGPT technologies maliciously. Below are some of the ways AI is being weaponized to attack organizations:



SECURING AI TECHNOLOGIES

Just as AI can be used to secure critical assets, in the wrong hands, it can also exploit them. Below are a few strategies organizations using ChatGPT or similar tools can adopt to protect their applications, data, and systems.



Policies and Procedures Specific to AI

Establishing guidelines for data usage, transparency, and accountability helps maintain the integrity and trustworthiness of AI systems. Microsoft created an AI security risk assessment framework to empower organizations to reliably audit, track, and improve the security of AI systems.[15] Whether your organization is using ChatGPT or an internally developed AI tool, it is important to develop clear governance documentation that can be shared across the organization.

Avoiding Fraudulent AI Applications

As ChatGPT grows in popularity and expands its user base, bad actors are using "spoofing" attacks, in which they masquerade as trusted AI software to trick users into providing sensitive information. Cyber criminals are also using AI to power social engineering attacks, such as phishing, vishing, and smishing. Educating staff about potential security risks, best practices, and proper usage of AI technologies can significantly enhance security.

Routine Data Hygiene

With enterprise ChatGPT, or any privately owned AI software, it is critical to routinely update the data that feeds the application. This will help reduce biases and ensure that the data is factual. Organizations should protect the data used to train and finetune AI models by implementing strong encryption methods, access controls, and data storage practices.

Access Control and Authentication Measures

Implementing strong authentication mechanisms ensures that only authorized individuals can access and modify the AI system. Multi-factor authentication (MFA), secure login procedures, and role-based access controls can be effective in preventing unauthorized access.

Continuous Monitoring and Frequent Updates

Continuous monitoring of AI systems allows for the detection of unusual behavior or malicious activities. Employing anomaly detection techniques, logging, and auditing mechanisms can help IT staff identify and respond to potential security breaches. Additionally, keeping the underlying software up to date with the latest security patches is essential to protect against cyber threats.

Regularly Scheduled Security Assessments

Conducting rigorous testing to identify and address potential vulnerabilities is crucial. Vulnerability assessments and penetration tests can help identify weaknesses in the AI system's defenses.

> One thing that humans and technology have in common is that they continue to evolve.[16] AI's timeline does not stop here. Innovations in AI will continue to transcend human expectations and bring creative solutions to a range of problems.

In addition to these tactics, technical protection measures on various levels should be applied. This includes securing the network via firewalls, endpoint protection, encryption, and data backup, and protecting hardware, operating systems, and software that thwart attacks and protect the input and output of the AI system.

Apart from network security, other general measures can also address AI-specific threats. One safeguard for the AI system development process is to mandate background checks on the developers. The organization should also document and protect important information cryptographically throughout the AI life cycle. This can include used data sets, pre-processing steps, pre-trained models, and the training procedure itself. Furthermore, conducting adversarial retraining on AI systems will enhance their resilience against cyber attacks.[17]

CONCLUSION

Automation for Everyone

It is undeniable that ChatGPT represents a breakthrough in enterprise AI. But, as our reliance on AI grows, striking a balance between innovation and security is essential to harnessing ChatGPT's potential without placing businesses and their customers at risk. Implementing robust data protection, authentication, regular updates, and monitoring can enhance the security of ChatGPT and other enterprise AI systems. In addition, ongoing research and collaboration within the AI community are crucial to staying ahead of emerging threats.

As we navigate the evolving AI landscape, it is imperative for developers, policymakers, and industry leaders to promote transparency, fairness, accountability, and security to fully harness the power of AI for a smarter and safer world. If you are considering using ChatGPT or an internally developed AI solution for your business, Securance can help you reach your cybersecurity and compliance goals. **Contact us** for a free consultation today.

ABOUT SECURANCE

Securance has more than two decades of experience helping organizations combat evolved cyber threats, build effective risk management programs, align with compliance standards, and increase operational efficiency. Our comprehensive approach integrates proven methodologies, dependable expertise, and each customer's unique requirements to maximize the benefits and long-term value of each assessment.

SOURCES

1. https://www.demandsage.com/chatgpt-statistics/
2. https://www.stylefactoryproductions.com/blog/chatgpt-statistics
3. https://www.gartner.com/en/articles/beyond-chatgpt-the-future-of-generative-ai-for-enterprises
4. https://research.aimultiple.com/chatgpt-for-business/
5. https://www.ibm.com/topics/artificial-intelligence
6. https://nordvpn.com/blog/what-is-chatgpt/
7. https://openai.com/blog/chatgpt
8. https://openaimaster.com/what-is-privategpt/
9. https://venturebeat.com/security/how-chatgpt-can-become-a-security-experts-copilot/
10. https://www.comptia.org/content/research/understanding-emerging-technology-artificial-intelligence
11. https://www.mailbutler.io/blog/email/ai-email-response/
12. https://www.forbes.com/sites/siladityaray/2023/05/02/samsung-bans-chatgpt-and-other-chatbots-for-employees-after-sensitive-code-leak/?sh=2e997e5f6078
13. https://hbr.org/2023/04/the-new-risks-chatgpt-poses-to-cybersecurity
14. https://www.forbes.com/sites/nishatalagala/2023/04/04/using-chatgpt-safely-the-legal-implications/
15. https://www.microsoft.com/en-us/security/blog/2021/12/09/best-practices-for-ai-security-risk-management/
16. https://verloop.io/blog/the-timeline-of-artificial-intelligence-from-the-1940s/
17. https://www.bsi.bund.de/SharedDocs/Downloads/EN/BSI/KI/Practical_AI-Security_Guide_2023.pdf?__blob=publicationFile&v=5

Proprietary ChatGPT: Enterprise AI Security Risks and Solutions for a Smarter World

SECURANCE CONSULTING
the advantage of insight

13916 Monroes Business Park, Suite 102, Tampa, FL 33635 • 877.578.0215
www.securanceconsulting.com