cybersixgill

FROM AI TO IQ

Transforming Cyber Defense with Generative AI

Introduction

Artificial intelligence – AI – is not a new concept. Dating back to 1956, it has taken several decades for AI be refined and become somewhat mainstream. Interest in the technology greatly expanded in the early 2000s as machine learning began to solve various problems in academia and industry.

Fast-forward to November 2022, when OpenAI launched the revolutionary Chat Generative Pretrained Transformer – better known as ChatGPT – the world has since been abuzz with the new chatbot and generative AI language tool. More than six months after the unveiling of ChatGPT, there are still many questions among organizations and individuals about what generative AI is and the benefits it offers. This is especially true in cybersecurity, where AI, specifically ChatGPT, is viewed as both a blessing and a curse. As much as AI enhances cyber defenders' ability to battle the ever-expanding threat landscape, cybercriminals also leverage the technology to streamline the development of new threats, mimic human language and behavior in launching attacks, and improve operational efficiencies.

This eBook examines the evolution of AI and generative AI and discusses their importance in cybersecurity. In particular, the report covers:

- An introduction to AI, LLM, ChatGPT, and BARD
- The importance of AI in cybersecurity
- How advanced AI is transforming cyber threat intelligence (CTI)
- Making CTI accessible and usable to all security maturity levels
- Selecting the right generative AI solution

Understanding AI: the similarities and differences of other technologies

AI is an area of computer science that focuses on creating and using intelligent machines that can perform tasks at speed and scale with human-like thinking – for example, learning, problem-solving, and decision-making. In the cybersecurity space, and with cyber threat intelligence (CTI) in particular, AI can analyze large amounts of data and identify patterns or anomalies that might indicate a potential threat.

Generative AI is a form of AI technology that can produce various types of content, including text, imagery, audio, and synthetic data. The excitement about generative AI largely stems from new tools (like ChatGPT) that can create high-quality text, graphics, and videos within seconds. Text-based Generative AI is built with Large Language Model (LLM) – an AI-based language model that can be used for natural language processing tasks. It works by analyzing large amounts of data, using that information to generate human-readable text or to gain insights about the data. In cybersecurity, LLMs can analyze and understand large amounts of data, such as social media posts, forum discussions, or threat reports, to identify potential threats.

GPT was a significant breakthrough in natural language processing because it was the first model to generate human-readable text with remarkable coherence and fluency. It achieved this by using a massive amount of training data and a complex neural network architecture, which allowed it to learn the patterns and structures of human language. ChatGPT was the first major iteration of GPT to gain widespread attention and adoption (which is still expanding). Similar GPT-based tools, such as Bard, which is Google's experimental AI chat service, have since been introduced. Bard is similar to ChatGPT in its ability to converse, but the main difference is that Google's service relies solely on information currently live on the web (no surprise!), while ChatGPT pulls its data from numerous sources – books and articles, blogs and online forums, Wikipedia, scientific journals, code repositories, and social media posts.

The impact that ChatGPT and similar generative AI tools can have on enterprises is still unfolding, but their potential is significant.

Gartner states, "ChatGPT — and other foundation models like it — is one of many hyper automation and AI innovations. It will form a part of architected solutions that automate, augment humans or machines, and autonomously execute business and IT processes. It will also likely be used to replace, recalibrate and redefine some of the activities and tasks included in various jobs. **ChatGPT is capable of the following:**

- Generating and helping to improve prose and code development
- Summarizing text
- Classifying content
- Answering questions
- Translating and converting language (including programming languages)[1]

Yet there are also limits. For example, ChatGPT currently only uses data through September 2021, so it can't be relied on for information about events taking place more recently. Additionally, it can't cite specific sources for its data, which means they may or may not be reliable, and therefore the knowledge it produces may not be accurate. ChatGPT also can't create images or perform math, and its efforts to ensure data privacy are inconsistent, so companies and their employees should exercise caution when using the tool.

1. Gartner (April 2023). Emerging Technology Horizon for Information Security, 2022

How AI Advances Cybersecurity and Threat Intelligence

There is no question that AI brings a wealth of benefits to cybersecurity. Instant report generation and simplifying complex threat intelligence to make it easier to use enables security teams to improve post-detection actions such as alert prioritization, augmented threat detection, playbook creation and incident response. It can also increase the efficiency, effectiveness and speed of alert generation and response, as well as improve analyst accuracy.

The ability of AI to enhance security outcomes at speed positions it to help fill current skills gaps within the sector. AI can increase accessibility to proactive attack detection without requiring a team of highly skilled security analysts. Further, given its ability to reduce manual work across multiple security areas, AI can enable an organization's existing team and resources to focus their time on more critical activities.

Not all generative AI solutions are the same

Currently generative AI is showing great promise – but with a caveat. Threat intelligence produces an abundance of data, but it can be overwhelming to security teams and is not always relevant to the organizations consuming it. In light of this, generative AI that relies on unique data provides the most significant value.

For the AI to give you the right answers, it must have access to the right data and intelligence, which is often hidden in underground sources. For example, if you want the AI tool to answer intelligence questions about activity on ransomware sites, it must access data from ransomware sites. If you want the AI tool to answer questions about initial access broker (IAB) markets, it must have access to IAB market intelligence. And if you want the AI to answer intel questions about threat actors sharing exploit codes in the underground, it must have access to vulnerability intelligence.

In short, many AI tools rely only on publicly available intelligence – which can quickly become outdated – or simply make up the answer when access to data is unavailable. Cybersecurity practitioners want AI to help them make an informed answer to critical questions, which means it must be fed from the broadest, deepest data lake of underground threat intelligence available.

Transforming CTI with Advanced AI

Cybersixgill is no stranger to AI. Since our inception in 2014, delivering valuable and impactful AI solutions has been in our company's DNA. Over the years we have launched unique automated models that score actors by their underground activities, allowing you to focus on the most nefarious threats, sophisticated profiling models that connect a threat actor's aliases and activity across the cyber underground and DVE Intelligence, which predicts a vulnerability's likelihood of exploitation so you can focus on the highest-risk CVEs.

With the AI market unfolding rapidly and new GPT-based solutions being introduced regularly, our vision for transforming threat intelligence with AI goes beyond a mere chatbot. We're on a mission to revolutionize threat intelligence, defying age-old axioms that only expert analysts can gain value from threat intelligence, shattering legacy notions of human vs. machine intelligence.

Introducing Cybersixgill IQ

Cybersixgill IQ is our new generative AI, representing a breakthrough in cyber threat intelligence. The solution draws from our unmatched, deep, dark web data and intelligence, as well as Open Source Intelligence (OSINT) and builds on our company origins firmly rooted in AI.

> *Generative AI can be a force multiplier, helping organizations derive value from threat intelligence. With Cybersixgill IQ, threat analysts and security professionals can now ask critical questions and get immediate, detailed answers, which can accelerate the value of CTI toward proactive investigations and understanding CVEs, exploits, IOCs, and TTPs. Failed threat intelligence programs are often the result of threat research outputs that are irrelevant to the organization. With generative AI capabilities such as Cybersixgill IQ, organizations can tailor threat intelligence and generate curated reports customized for the various constituents consuming them, including CISOs, SOC engineers, business managers, and everything in between.*
>
> Jon Oltsik, distinguished Analyst, Fellow with Enterprise Strategy Group and the founder of the firm's cybersecurity service

While other generative AI solutions rely on simple integrations with ChatGPT, Cybersixgill IQ leverages AI across capabilities, supporting every step of the intelligence process. The solution simplifies access to CTI, making it easier to answer complex intelligence-related questions with readily available, actionable insights. Supplementing and enhancing our deep, dark web threat intelligence through our intuitive portal or API, the generative AI tool transforms access to CTI and our expert insights in three key ways:

Intel analysis

Turns tactical, raw intelligence, previously only accessible to experienced CTI analysts, into more easily understood, context-and-insight-rich summaries that are usable and actionable for every cybersecurity professional. For example, the solution summarizes and auto-distills insights about victims of initial access broker markets and ransomware sites.

Intel Generation

Delivers finished intelligence in seconds, using AI to automatically and dynamically generate on-the-fly, refined, finished intelligence customized to each organization's industry, geography, user persona, and business needs. For example, users can ask Cybersixgill IQ to generate a report about ransomware attacks in 2023 targeting North American retailers.

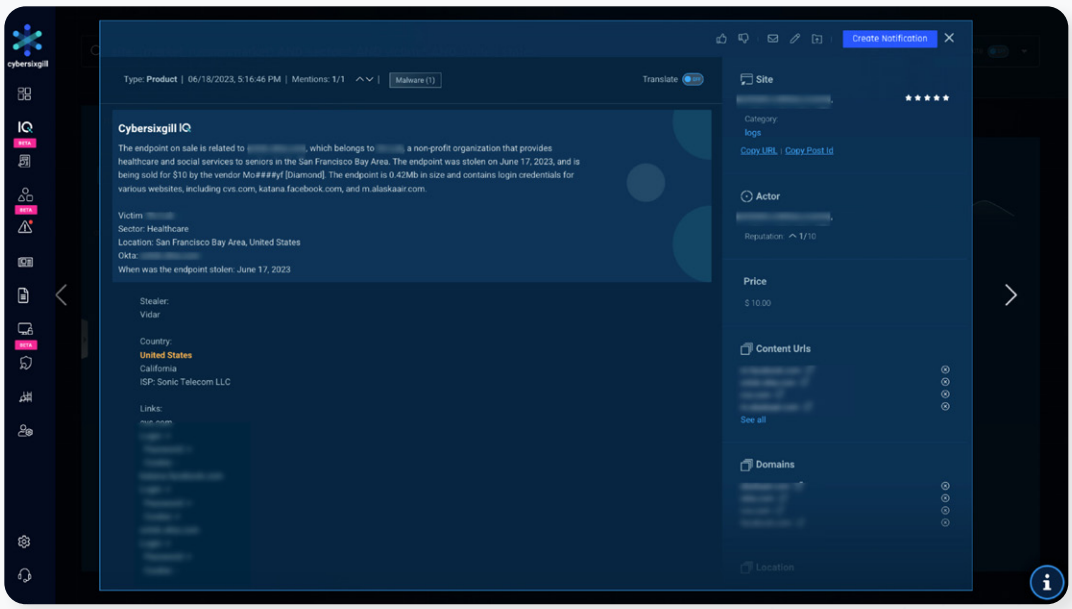
Intel Experience

Transforms how users interact with CTI. Cybersixgill IQ's insights are delivered through a threat intelligence chatbot, a context-specific assistant, or across the Cybersixgill portal. For example, users can ask the chatbot to analyze a CVE's likelihood of exploitation based on Cybersixgill's proprietary vulnerability intelligence. In essence, the solution enables users to shift their focus and energy from comprehending complex data to quickly answering critical intelligence questions.

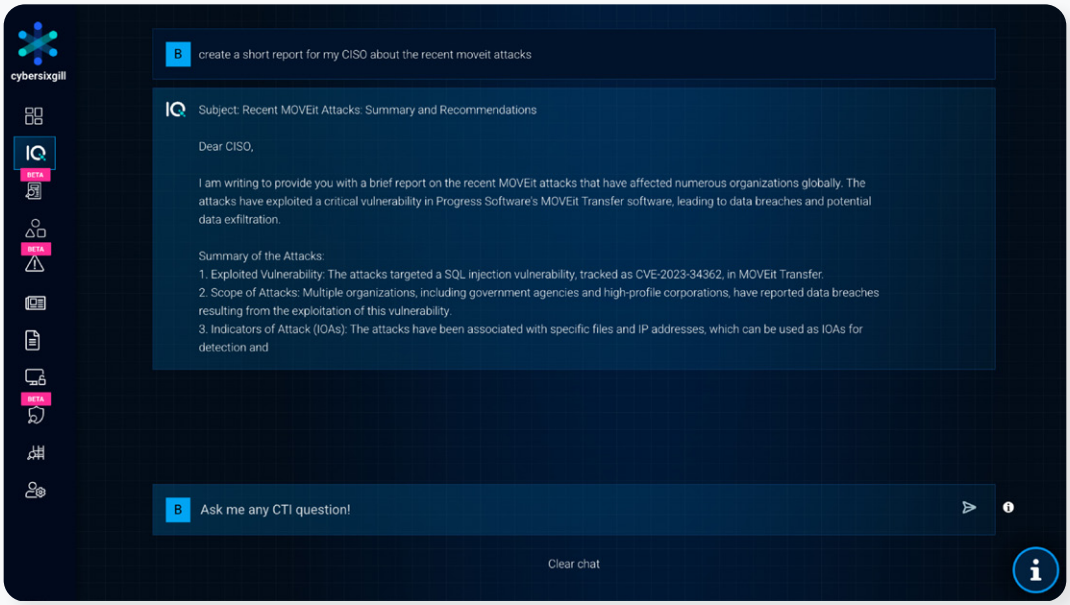
Democratizing CTI for cybersecurity defenders

Cybersixgill IQ leverages state-of-the-art generative AI technologies to address an extensive set of business use cases, improving efficiencies with limited resources, and democratizing CTI for organizations of all sizes and security maturity levels. The tool also serves a range of users across job titles and skill levels, such as:

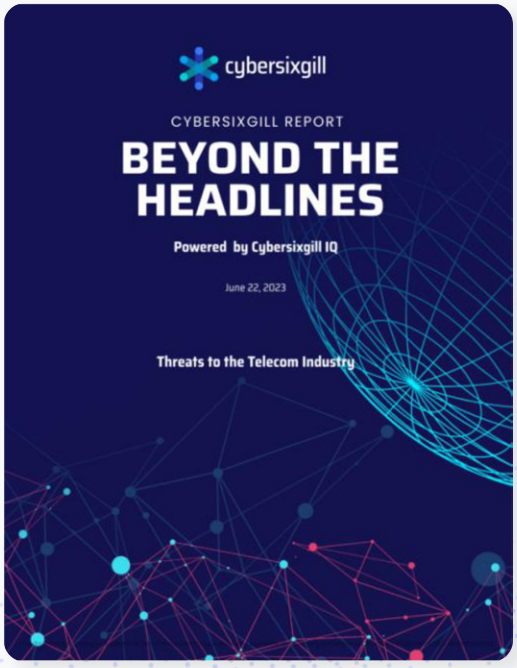
- Beginner analysts can close the skills gap with senior analysts, enhancing their CTI capabilities 10X.

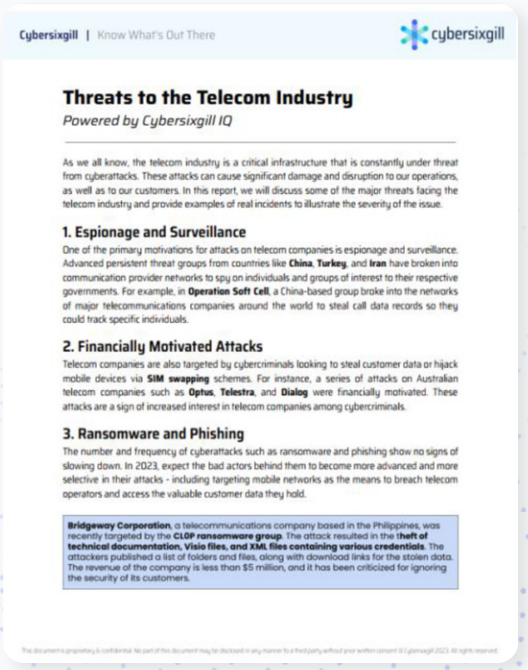


- Advanced users can reduce the amount of tedious work, focus on asking the right intelligence questions, and perform advanced analysis that only humans can.

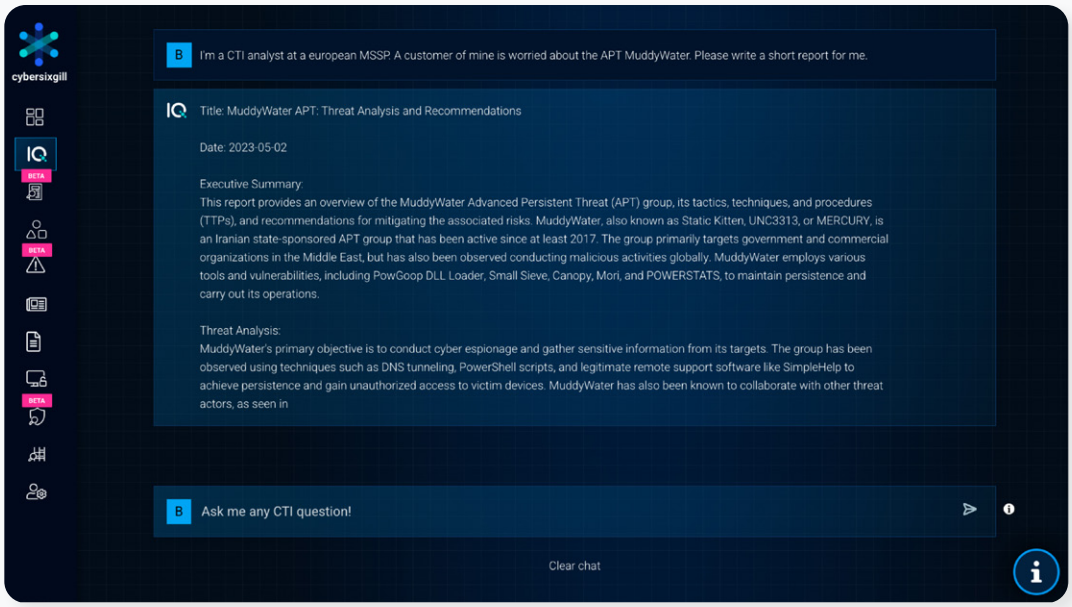


- CISOs and senior executives receive fully digested reporting dynamically fitting their specific needs.





- Managed Security Service Providers (MSSPs) can quickly and efficiently scale the delivery of intelligence for customers, increasing their ability to service more organizations while maintaining a high-quality standard.



Market Selection: Finding the Right AI Tool

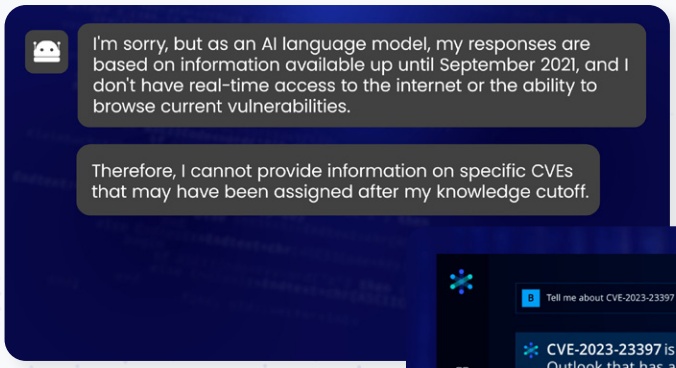
With the advances that generative AI enables, understandably it is gaining a firm foothold in the cybersecurity industry. While legitimate concerns exist around data privacy, misuse of information, and bias, adoption will only continue over time. Anyone who stays on the sidelines risks falling behind.

To help users find the right solution for their needs, we have compiled a list of important capabilities and attributes you should look for, while highlighting some gimmicks and marketing hype you should also be aware of.

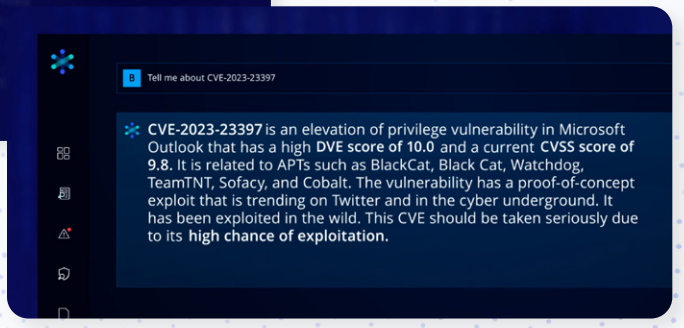
1. Breadth and depth of data

As stated in this e-guide, unique data is critical when relying on generative AI for credible answers and information. Off-the-shelf and open-source AI solutions are only as good as the data they access, and most available solutions have access to a limited set of sources. For example, if you ask chatGPT a question about something that happened in a deep web forum or on a dark web market, the response you receive will either be inaccurate or left blank, given its lack of access to this intel.

Cybersixgill IQ intelligently interprets customer inquiries, delivering data and insights that precisely align with their required use cases in the format you need – be it a concise threat overview summary for your CEO, vulnerability exposure analysis for MSSPs, or a comprehensive forensic incident report for detection and response teams.



ChatGPT response to a request for information about CVE-2023-23397

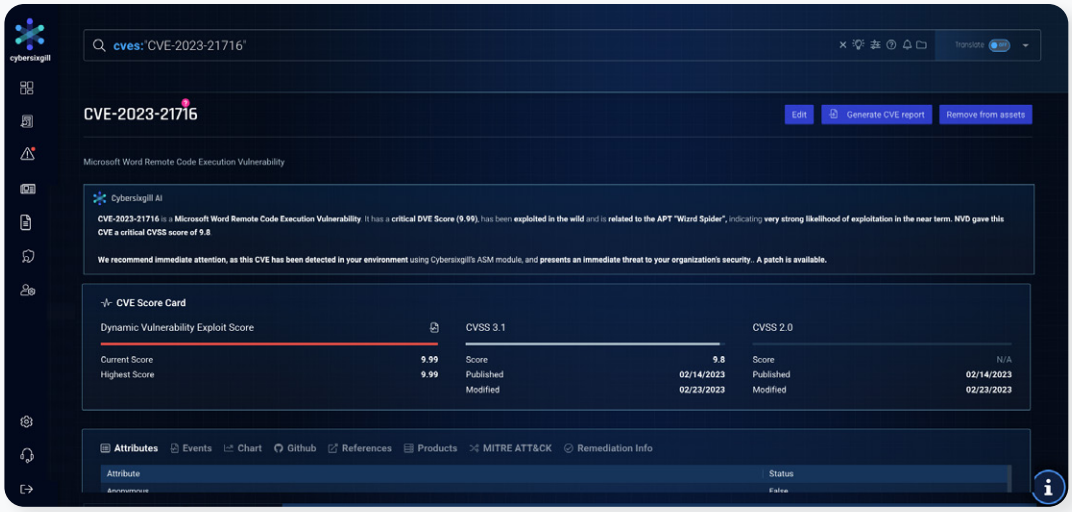


Cybersixgill IQ response to a request for information about CVE-2023-23397

2. Proven history and credibility with AI

Not all AI solutions are alike, and not all CTI vendors have solutions that are rooted in AI.

Cybersixgill has always invested in AI and the automated processing and enrichment of our data. We've implemented machine learning and deep learning for the past several years, as evidenced by products such as DVE Intelligence, which leverages real-time NLP-driven analysis of underground chatter to predict a CVE's likelihood of exploitation in the near term.



3. Beyond chat

Most generative AI solutions simply offer a chat feature, which can be helpful in some instances but doesn't offer the level of actionable information you need to hasten critical decision-making.

In contrast, we've embedded AI across the Cybersixgill IQ solution - from human-readable, automated analysis of intelligence items or the immediate generation of high-quality intelligence reports, to an AI analyst assistant that follows your work, providing vital insights at every stage.

4. Protecting the privacy of your data

As previously mentioned, concerns about data privacy in existing AI solutions are well-founded and should be taken seriously when choosing a generative AI tool.

At Cybersixgill, we take legal and ethical implications seriously and have implemented measures to ensure that our customers and their data's privacy and security are upheld. As we enter the new age of AI, we're implementing our solutions with a cautious, security-first approach and do not send customer data to services like ChatGPT.

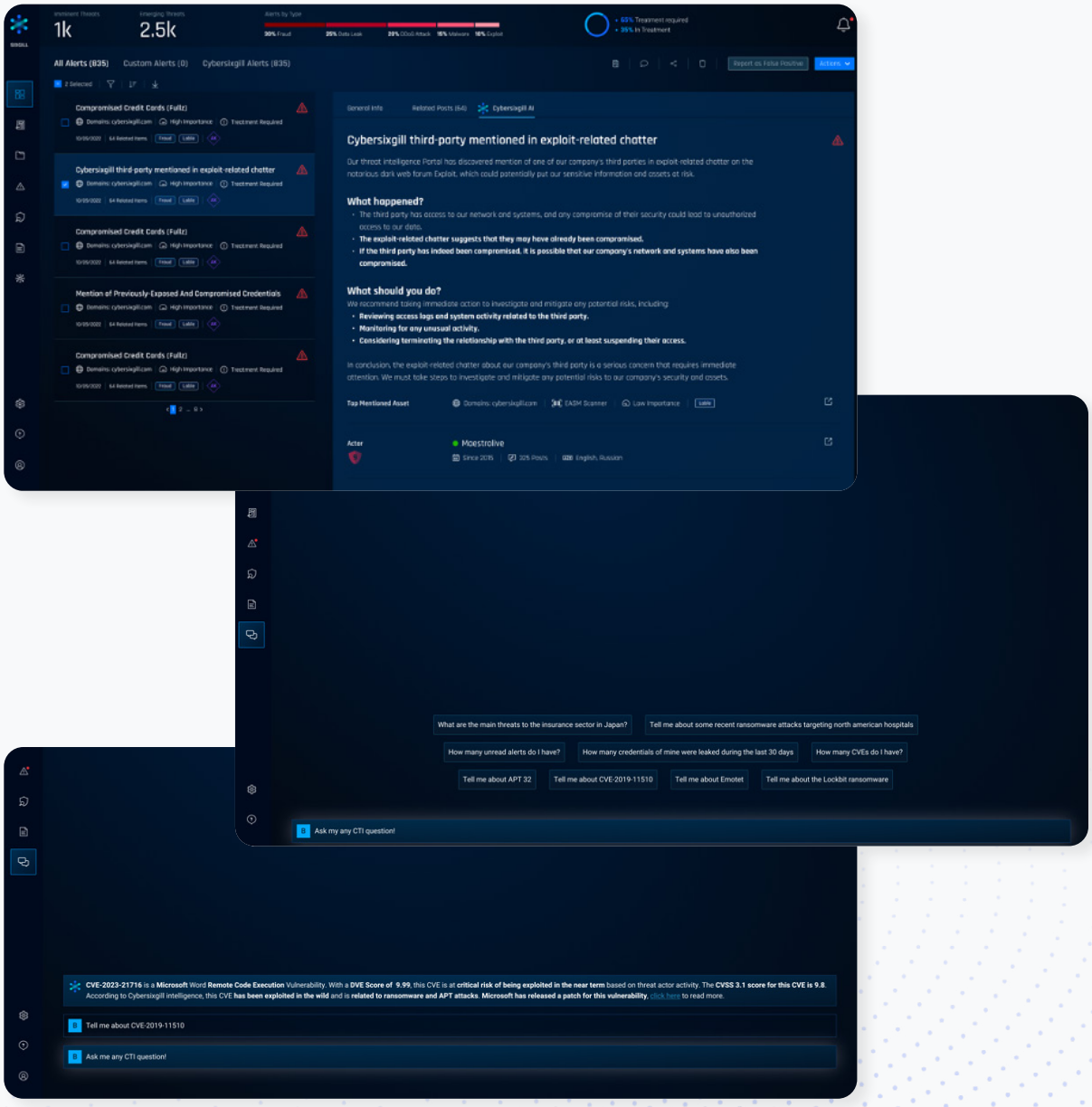
5. Protection against AI hallucinations

Off-the-shelf LLMs like GPT and Bard may occasionally generate "fake" or hallucinated content.

Cybersixgill IQ is designed to mitigate this issue in several ways, such as:

- Processing model answers before delivering an output to the end user to avoid "common" hallucinations. This includes building a proprietary repository of common GPT mistakes and increasing fidelity over time.
- Excluding answers if the AI is unsure about the result.
- Focusing the model's querying of the data using scoped data access and prompt engineering. (Prompt engineering is the process of designing and refining prompts to achieve specific goals, such as generating content for marketing campaigns or identifying relevant information in social media posts.)
- Robust feedback mechanisms to engage in fast feedback loops with users and detect and mitigate incorrect AI-generated content.
- Manual review of the model using our in-house data analyst team and continuous improvements to review and refine the model manually.
- Training our proprietary transformer (GPT) CTI model to fit our specific domain, unique data, and customers' use cases. This enables us to leverage the benefits of large language models and transformer technology to provide the distinct context of threat intelligence. (Transformer is an AI-based neural network architecture that can be used for natural language processing tasks. It's used in models such as GPT and LLM to generate or analyze human-like text.)

In the images below, you can see examples of the outputs that Cybersixgill IQ offers for specific questions, compared to the outputs you'd receive from other chatbots.



How Cybersixgill protects the privacy of your data

Generative AI is a promising field with exciting potential. There are risks involved with relying on third-party vendors to handle sensitive customer information. We take these risks seriously. In addition to the Data Processing Addendum (DPA), we have implemented further measures to ensure your data security and privacy:

Minimizing Data Transfer

By employing efficient data reduction strategies and making smart use of local resources, we ensure that only the most essential, non-sensitive information is shared.

Masking Sensitive Data

Before sharing data with OpenAI or any other third party, we replace actual data with randomized characters or other "noise," ensuring that while the data structure remains intact for analysis, the sensitive information within is well-secured.

Sending Metadata Only

In certain scenarios, we only send metadata to OpenAI. Metadata is essentially the 'data about the data' – it doesn't include the actual content but contains details about it.

Using Differential Privacy

This is a technique where we publicly share information about a dataset by describing group patterns within the dataset while withholding individual-specific information. This ensures that individual privacy risk is mathematically bound, even amidst external information.

Local Processing

We prioritize local data processing to limit the amount of data transferred over the internet. This may involve extracting features from the data, converting it into lower-dimensional representations, or using local models to anonymize it before it's sent to OpenAI.

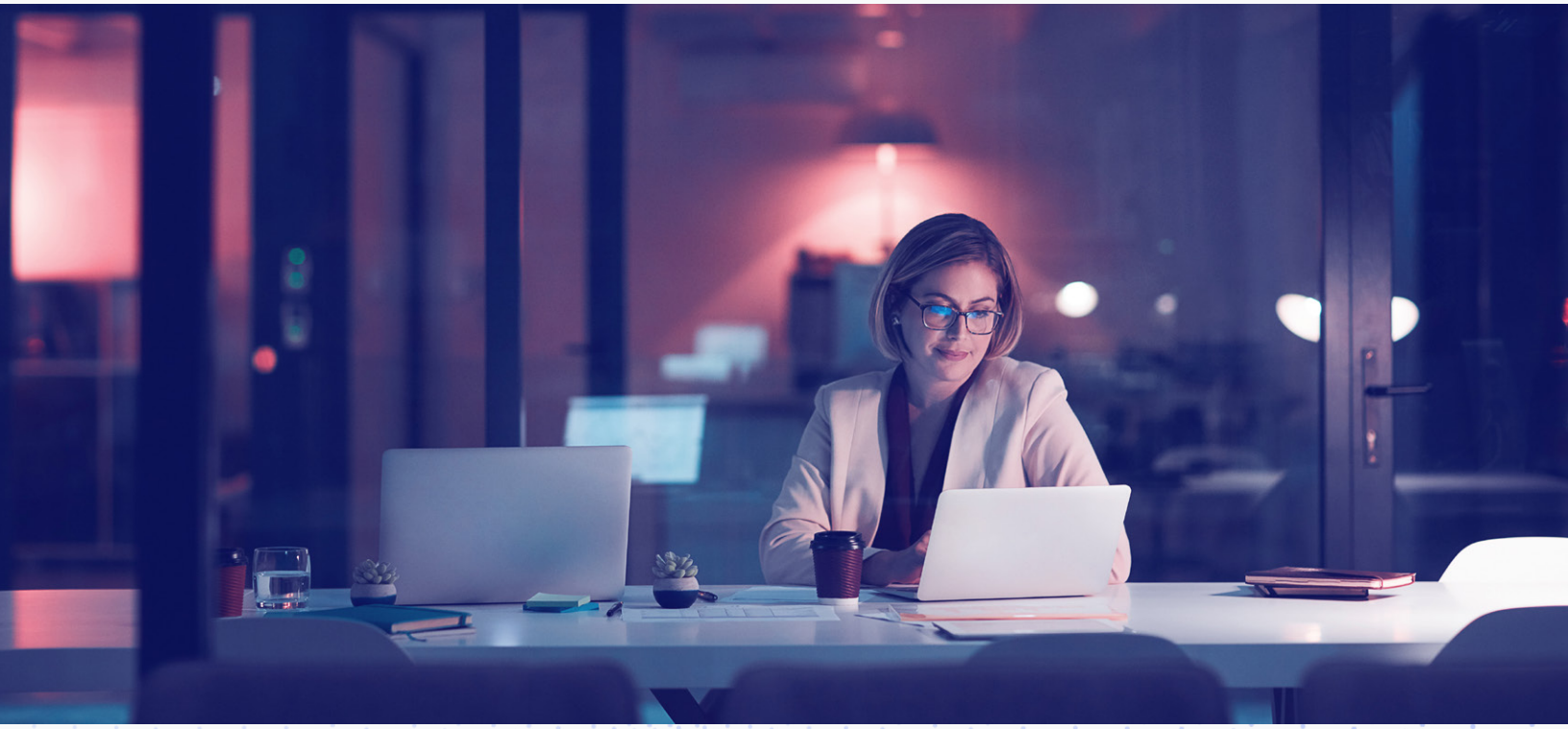
Developing Our Proprietary Models

To further tighten our data security measures, we invest in developing our proprietary machine learning models. These models can be trained on our sensitive data, but without the data ever leaving our secure servers. It allows us to maintain control and ownership of our data and the insights derived from it.

Summary

This new era of advanced AI allows us to make CTI accessible, practical, and usable to more cybersecurity professionals, regardless of security maturity. As we transform CTI with advanced AI, more skilled analysts can do less tedious CTI work and focus on more complex challenges, as they can ask more complicated intelligence questions. Most importantly, communities previously excluded from using CTI due to lack of time, money, and resources, can now reap the benefits of fully processed, automatically-generated intelligence.

With the right data pulled from underground sources that aren't broadly accessible, AI can significantly improve your cybersecurity efforts and defensive posture. While AI and generative AI technologies are still evolving, recent advances are already delivering significant benefits to organizations taking the plunge. Cybersixgill IQ is on target to redefine the CTI landscape, providing unprecedented access and insights to help cybersecurity defenders in every industry and every level of experience.



About Cybersixgill

Cybersixgill continuously collects and exposes the earliest indications of risk by threat actors moments after they surface on the clear, deep, and dark web. Our proprietary algorithms extract data from a wide range of sources, including content from limited-access deep and dark web forums, underground markets, invite-only messaging groups, code repositories, paste sites, and clear web platforms, as well as an unparalleled archive of indexed, searchable historical data from as early as the 1990s. This data is processed, correlated, and enriched with automation and advanced AI to create profiles and patterns of malicious threat actors and their peer networks, and deliver critical insight into the nature, source and context of each threat.

Our extensive body of threat intelligence data can be consumed through various solution offerings and integrations, each addressing critical customer pain points and use cases. These solutions are scalable, searchable and seamlessly integrated into existing security stacks, quickly arming enterprises, government and MSSPs alike with accurate, relevant and actionable insights to proactively block threats before they materialize into attacks.

Learn more at www.cybersixgill.com

Follow us

cybersixgill