How AI Helps Improve Your Security Posture

Written for KAMIND IT, Inc.

By Ravi Das

Contents

How AI Helps Improve Your Security Posture

Introduction

The world of Cybersecurity is known for its plethora of techno jargon, but there are now two that are being bandied about. They are those of Artificial Intelligence and Machine Learning. Truthfully, there is really nothing new about these two practices. They have been around since the mid-1950s, and have evolved since then. But the craze in which they are both now in today has been largely fueled by the dawn of ChatGPT, the latest AI/ML tool from OpenAI.

With it, you can ask the platform anything you want, and it will come up with an answer for you. In fact, an end user can even create an entire document from it, just by merely typing in a simply query into the search tool. Because of this level of sophistication, AI and ML have now started a widespread fear of its potential uses. For example, there is huge angst if AI and ML will replace jobs, or even be used for surveillance purposes without out knowledge.

The good news here is that all of these fears are mostly just that – myths. What is fueling all of this is that there is the belief that we will be able to replicate the thought and reasoning powers of the human brain. But the bottom line is that we will never understand the human brain, unlike some of the other organs in our system.

It is important to keep in mind that the human brain is a very complex organism – at best, we will only come close to understanding 0.5% of it. But there are areas in which both AI and ML will work very well – one of these is in Cybersecurity, which is the focal point of this whitepaper. But first, we start off with a review of what both AI and ML are all about.

What Is Artificial Intelligence?

When one thinks of the term "Artificial Intelligence", very often, an image of the human brain is conjured up. To a large extent, this is a good graphical representation of what it is. The basic idea of this is to mimic the thought and behavioral process of the human brain, to learn, and to apply these "experiences" towards discovering hidden trends and predicting the future. Specifically, artificial intelligence can be defined as follows:

"The term artificial intelligence (AI) refers to computing systems that perform tasks normally considered within the realm of human decision making. These software-driven systems and intelligent agents incorporate advanced data analytics and Big Data applications. AI systems leverage this knowledge repository to make decisions and take actions that approximate cognitive functions, including learning and problem solving."

Artificial intelligence (AI) also makes it possible for machines to learn from experience, adjust to new inputs and perform human-like tasks."

(SOURCE: 1 and 2).

As one can see from the above definition, the potential that Artificial Intelligence has in Cybersecurity is quite enormous. For example, many IT security teams of today use many kinds of tools from various kinds of vendors in order to beef up their lines of defenses.

As a result, they are completely inundated and flooded with all sorts of information and data coming in from all sources. Because of this, there are many false positive warnings and messages that are created, thus putting a huge extra burden for the IT security team to filter through.

But with the use of Artificial Intelligence, it can use the concepts of data warehousing in order to instantaneously filter through all of this and provide warnings and messages that are truly indicative of an impending threat vector. In this regard, Artificial Intelligence has also started to make its splash as a task automation tool for both Penetration Testing and Threat Hunting Teams.

As a result, these teams can now devote more time and attention to unearthing all unknown gaps and vulnerabilities that are in existence from within an IT Infrastructure and providing timely solutions for the client on the re-mediative actions that they need to take.

Another area where Artificial Intelligence is in the area of staff augmentation. At the present time, there is a severe worker shortage in the Cybersecurity Industry, as there are currently some 3.5 million jobs that still remain unfilled.

(SOURCE: 3).

In this regard, tools can be used to help predict what the future Cyber Threat Landscape will look like, and even provide recommendations to the IT Security Staff as to prepare for it how best. Typically, this would be the role for the Cybersecurity Analyst, but until enough positions are filled in this aspect, Artificial Intelligence will literally become the "virtual" version of it.

In fact, Artificial Intelligence (AI) can be further subdivided into two main types of classification, which are as follows:

1) Type 1:
   This consists of the following:

   - Weak, or "Narrow" based Artificial Intelligence:
     In this scenario, the AI system is focused on just accomplishing one specific task or goal. This can be considered to be a very primitive form of it, as it can only do very basic tasks, such as playing against a competitor in a basic Chess game. In this situation, all the rules and probable scenarios must be fed manually into the AI system before it can engage a true competitor. In other words, it cannot learn on its own, each and every time that a new game is played, all the rules and outcomes have to be fed into it each and every time.

   - Strong Artificial Intelligence:

This is the kind of AI that is used most typically in Cybersecurity today. For example, based upon the data and intelligence feeds that are fed into it, this kind of system can literally learn all on its own from past observations, and use that knowledge in order to make informative decisions for the future. In other words, every effort is made so that it can emulate human thought and decision-making process as much as possible. The key differentiator here is that there is almost no human intervention needed for these kinds of AI systems; the only time this is really ever needed is if new information and data feeds have to be inserted.

2) Type 2:
These kinds of Artificial Intelligence systems are based upon the functionalities that they possess or are anticipated to in the future. These include the following:

- Reactive Machines:
This is considered to be the most rudimentary, or basic form of an AI system, and in fact is more like a Type 1, as just reviewed. In particular, their memory is very limited, and these kinds of systems can only store very limited amounts of information and data and cannot make future decisions or predictions on their own, without some sort of human intervention.

- Limited Memory:
These are the Artificial Intelligence systems that can “learn” from past examples and use those to make decisions for future events. In fact, this is the classification that is more representative of AI systems that are in existence today, and that are used in the Cybersecurity Industry today.

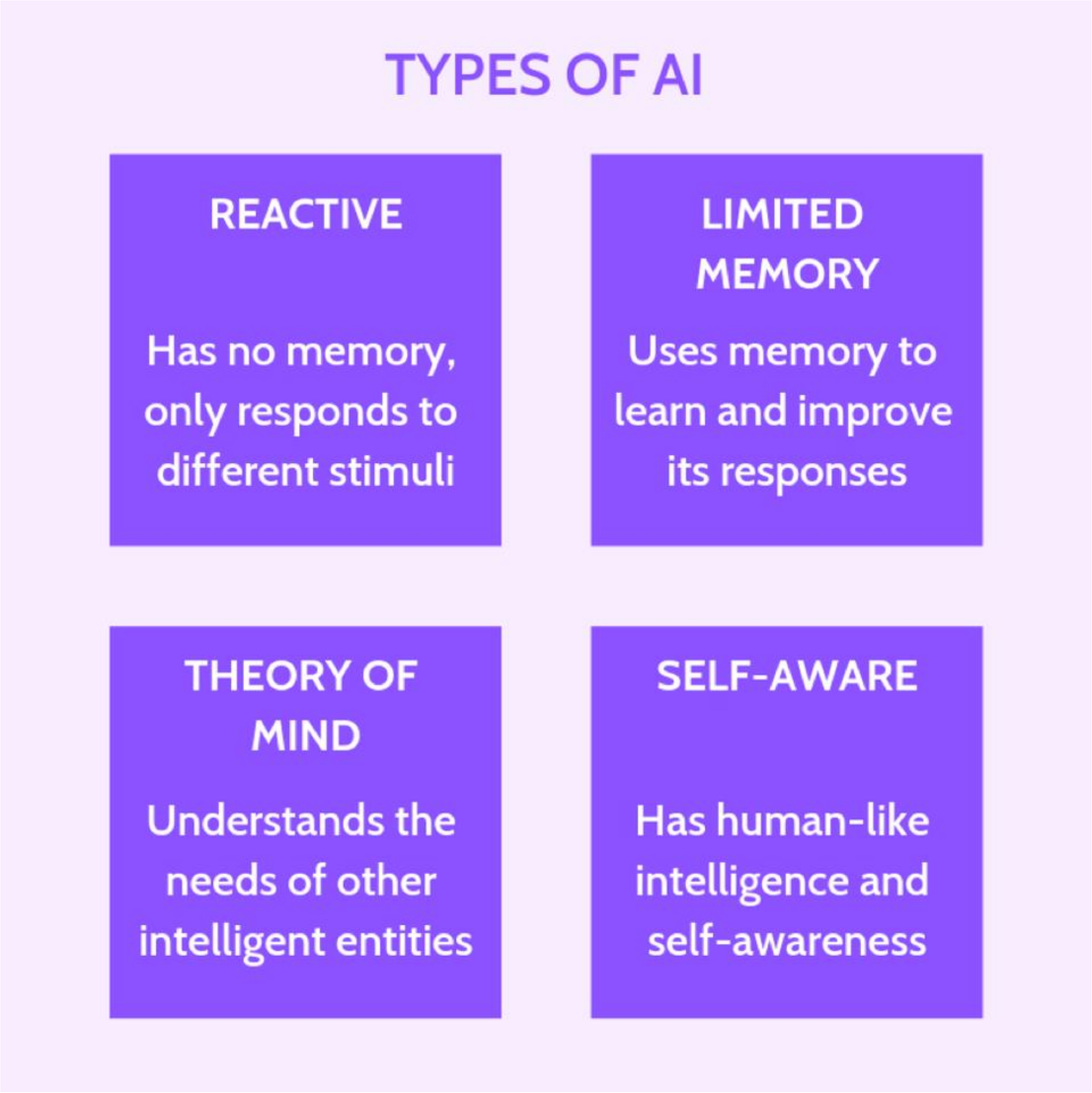
- The Theory of the Mind:
The main purpose of this kind of Artificial Intelligence system is to “... emotion, belief, thoughts, expectations and be able to interact socially.”

  (SOURCE: 4).

  Although this has a very long way to go until it can even come remotely close to achieving the above, we are already seeing some very basic forms of this starting to take place. The best examples of this are the Virtual Personal Assistants, namely those of Alexa, Siri, and Cortana. These kinds of applications try to learn our thought and decision-making profile so that recommendations and directions can be provided based upon previous actions.

- Self-Awareness:
This would be the ultimate goal of any Artificial Intelligence system. This entity would be very much like a being and even act like one. The best, illustrative example of this is the character Data from the TV series, “Star Trek: The Next Generation”.

These two classifications are illustrated below:



TYPES OF AI

REACTIVE

Has no memory, only responds to different stimuli

LIMITED MEMORY

Uses memory to learn and improve its responses

THEORY OF MIND

Understands the needs of other intelligent entities

SELF-AWARE

Has human-like intelligence and self-awareness

(SOURCE: 4).

A Definition of Machine Learning

Machine Learning, also known as “ML”, can be specifically defined as follows:

“Machine learning is an application of artificial intelligence (AI) that provides systems with the ability to automatically learn and improve from experience without being explicitly programmed. Machine learning focuses on the development of computer programs that can access data and use it to learn for themselves.

The process of learning begins with observations or data, such as examples, direct experience, or instruction, in order to look for patterns in data and make better decisions in the future based on the examples that we provide. The primary aim is to allow the computers learn automatically without human intervention or assistance and adjust actions accordingly."

(SOURCE: 5)

To break this down, ML is a subset of a much larger field known as "Artificial Intelligence". One of the primary goals of Machine Learning is to allow a Cybersecurity system to learn from a prearranged set of information and data, without any human assistance. From what it learns, this system is then able to project what the future holds in terms of threat vectors. Another objective of Machine Learning is to help filter out what alerts and warnings are for real, and which are not.

But keep in mind, in a very simplistic view, Machine Learning takes on more a "Garbage In, Garbage Out" methodology. In other words, the system is only as good as the information and data that is fed into it. Thus, in order to keep this system optimized on a real time basis, it must keep receiving and digesting various sorts of data sets 24 X 7 X 365. A cardinal rule of thumb in this instance is that the more intelligence feeds are used, the better, as this will lead the Cybersecurity system to discover a plethora of unhidden trends. As a result, this will make that much more robust in helping to combat Cyberattacks.

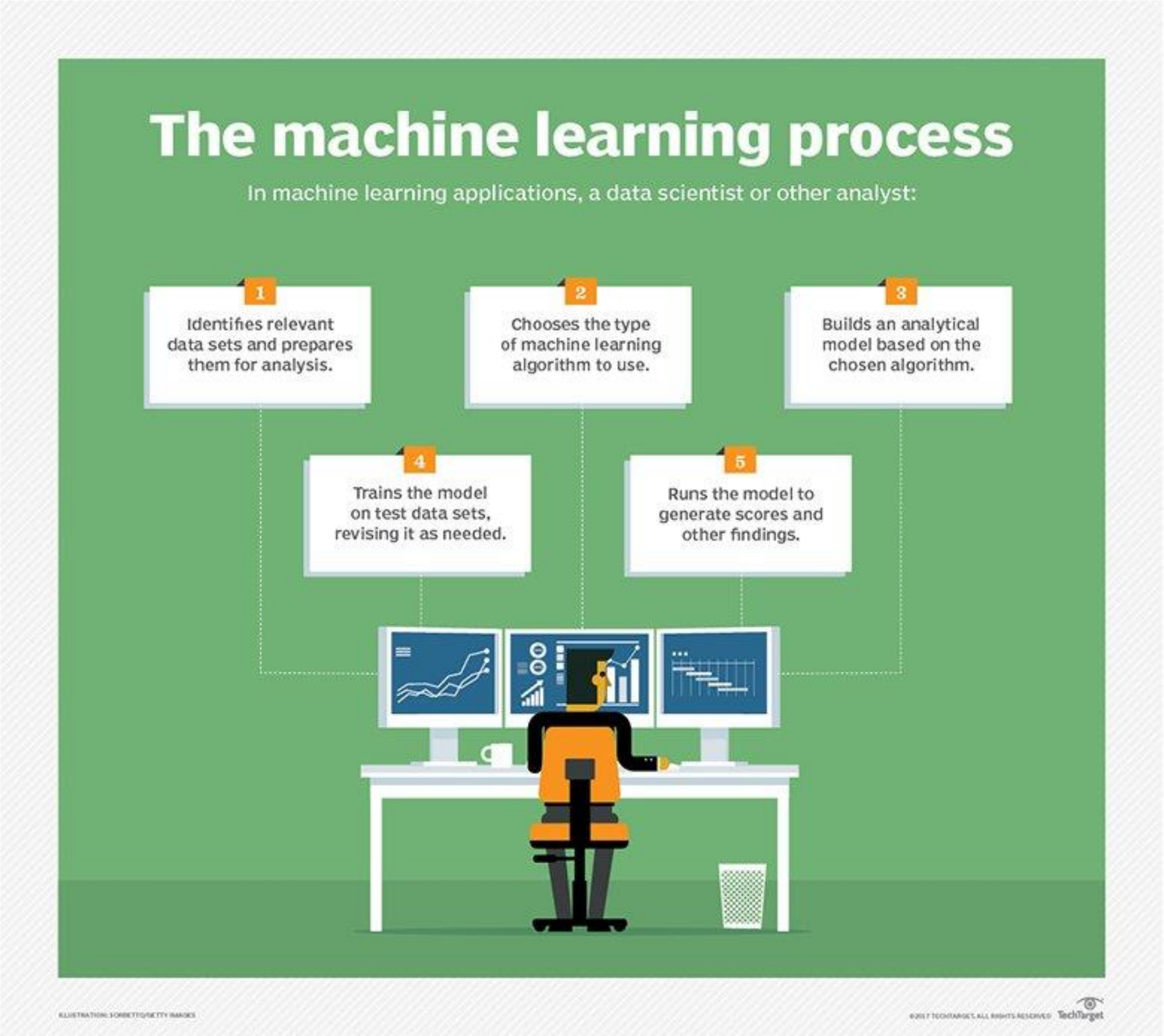
How Machine Learning Works

Machine Learning works in two different formats, depending upon the security requirements of the business or corporation:

1) A Supervised Approach:
   In this scenario, the Cybersecurity system that used Machine Learning requires human intervention to varying degrees. For example, a highly skilled data analyst is often required in order to select the inputs and feed them. Also, the desired output must be programmed into it as well, so that it can provide the appropriate answers that are desired.

2) An Unsupervised Approach:
   This is where the Cybersecurity system can simply be fed the many intelligence feeds that it requires (and as mentioned, the more of this the better). From there, it will learn the various trends in the datasets, and even discover hidden or hard to notice trends and statistical correlations as well. These kinds of systems do not need to be programmed for what the desired outcome will be. In order to accomplish these sorts of tasks, very sophisticated algorithms known as "Neural Networks" and "Deep Learning" are very often used.

Both approaches are illustrated in the diagram below:



The machine learning process

In machine learning applications, a data scientist or other analyst:

1 Identifies relevant data sets and prepares them for analysis.

2 Chooses the type of machine learning algorithm to use.

3 Builds an analytical model based on the chosen algorithm.

4 Trains the model on test data sets, revising it as needed.

5 Runs the model to generate scores and other findings.

ILLUSTRATION: SORBETTO/GETTY IMAGES

©2017 TECHTARGET. ALL RIGHTS RESERVED TechTarget

(SOURCE: 6).

How AI & ML Can Improve Your Cybersecurity Posture

Both AI and ML can serve a wide variety of both needs and applications   Examples of these include the following:

Proactive Monitoring & Remediation – Reducing Alert Fatigue

There is one common denominator in the Cyber world of today: It never sleeps. The Cyberattacker, wherever they might be located in terms of geography, are always coming up with ways in how to craft together the newest threat variants, which can cause the maximum amount of damage. But it is important to keep in mind here that the Cyberattacker will never come up with a total, 100% threat variant from scratch. Rather, they will simply try to build a better mouse trap, in which they will simply take an existing variant, and just tweak it enough so that it will go undetected initially, and inflicting more pain to the unsuspecting victim than ever before.

Because of this, and the nonstop barrage of attacks, the IT Security team of today has to be on their constant guard at all times. While in theory this is possible, but the reality dictates the opposite. Humans simply cannot stay up for days at a time, to simply watch the threat board of what is going on, and trigger any alerts and warnings that come through. CISOs have tried to do this, but now they are starting to realize that their own employees have their breaking points, given the high burnout rate.

As a result, members of the IT Security team become too tired and cannot even discern what is real and what is not. This phenomenon has a technical name to it, and it is called "Alert Fatigue". It can be defined as follows:

"Alert fatigue is a phenomenon that occurs when cybersecurity professionals are inundated with such a high volume of security alerts that it leads to a diminished ability to react effectively to and investigate real threats.

Alert fatigue typically manifests from not filtering or prioritizing alert-triggering issues as well as unmanaged incoming notifications. In many cases, it's a combination of these reasons."

(SOURCE: 7).

So as you can see, Alert Fatigue can have very severe consequences, with the extreme being a total security breach impacting an entire business or even industry. But this is where both AI and ML can help. Once the model has been created, the system needs to start learning. This can be done by having it ingest a large amount of network traffic data, and from there, having it learn what behaviors are normal and what is out of the baseline profile.

From here, it can then send to the IT Security team those alerts and warnings that are real. This totally eliminates the need for manual inspection and review. The biggest advantage here is that they can be triaged in just a matter of seconds rather than hours with the traditional methods. As a result, threat variants can thwarted off almost immediately before they do indeed penetrate the IT and Network infrastructure of a business.

The AI/ML system can even do the following:

- Tier all alerts and warnings based upon a prioritization schedule.
- Consolidate all similar or identical warnings into one.
- Assign the level of action that needs to take place.
- Keep reviewing the alerts and warnings, and learn from them, in order to keep the algorithms refined and optimized.

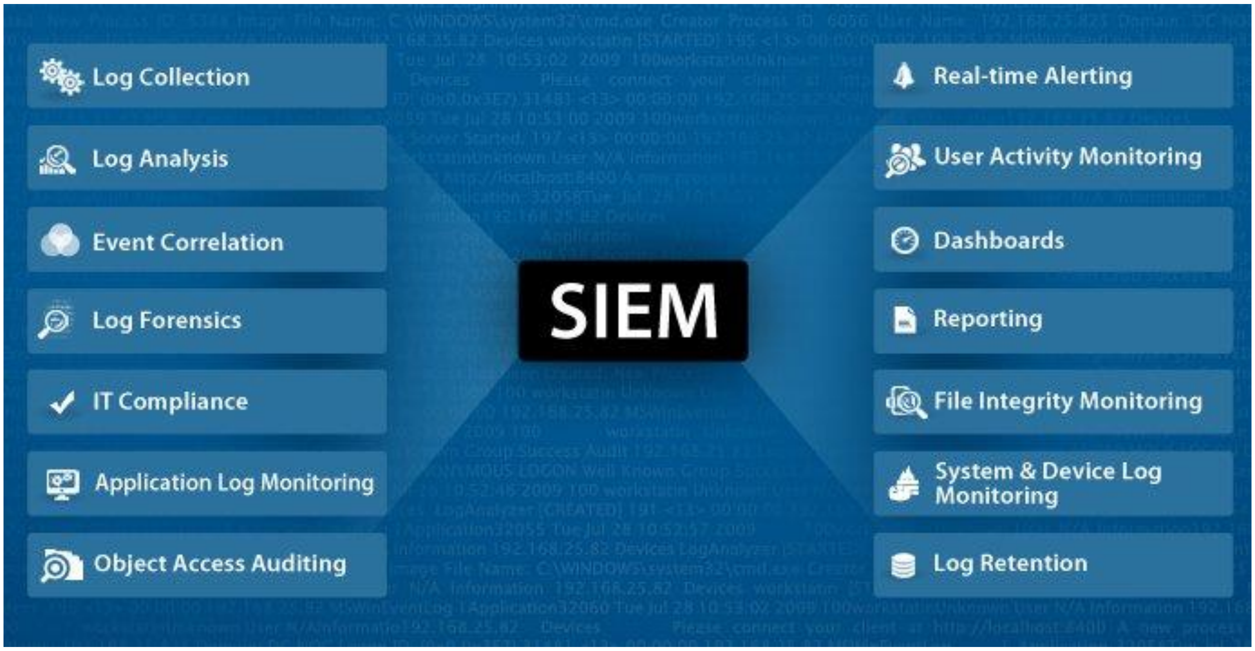
For this kind of application, a tool that is needed is what is known as a "Security Information & Event Management" system, or simply known as a "SIEM". It can be defined as follows:

“SIEM systems help enterprise security teams detect user behavior anomalies and use artificial intelligence (AI) to automate many of the manual processes associated with threat detection and incident response.”

(SOURCE: 8).

In simpler terms, a SIEM is just like a central dashboard where the IT Security team can see all legitimate warnings and alerts in one view, fed by the AI/ML system.

An example of a SIEM is illustrated below:



(SOURCE: 9).

Proactive Monitoring & Remediation- Automated and Intelligent Cyber Defense

Most businesses of today use two types of major tools to help not only fight off the threat variants, but to see where any hidden gaps and vulnerabilities may exist at. These are what are known as Vulnerability Scanning and Penetration Testing.

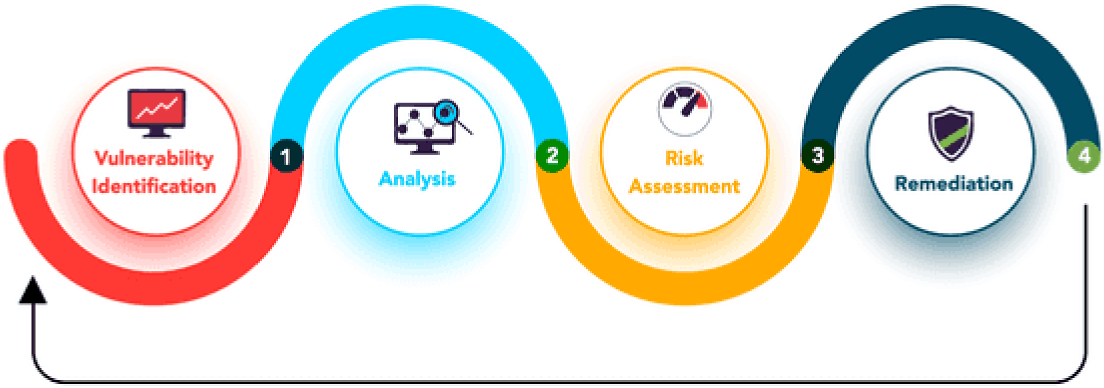
The former can be defined as:

“Vulnerability scanning is the process of inspecting and reporting potential vulnerabilities and security loopholes on a computer, network, web application or other device, including switches, routers, firewalls and wireless access points.”

(SOURCE: 10).

Simply put, Vulnerability Scanning is just a passive based one, which looks primarily at your Network Infrastructure. One of the key areas it looks for are open ports, which is often a forgotten about item by the IT Security team, causing a grave security threat.

An example of Vulnerability Scanning is illustrated below:



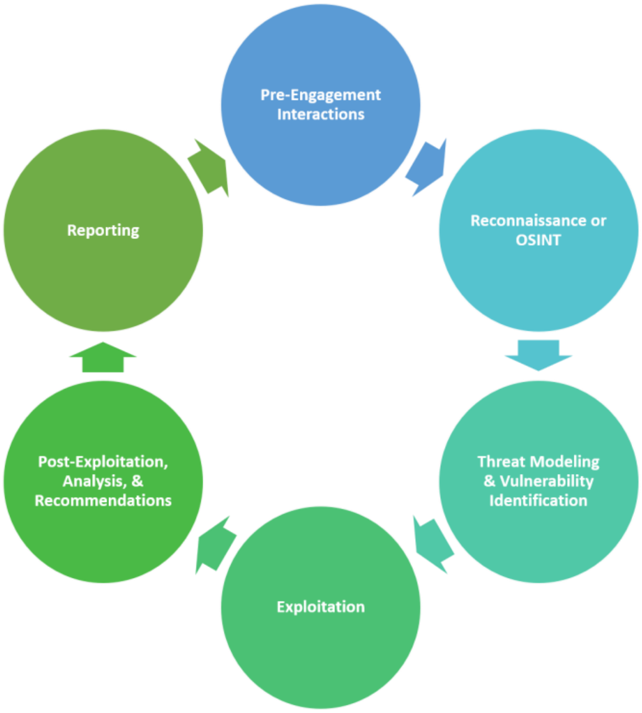
(SOURCE: 11).

The next major methodology that is considered to be a step up from Vulnerability Scanning is that of Penetration Testing. It can be defined as follows:

"Penetration testing, also called pen testing, is a cyberattack simulation launched on your computer system. The simulation helps discover points of exploitation and test IT breach security."

(SOURCE: 12).

As you can see, Penetration Testing is far superior to Vulnerability Scanning. The people that are involved in doing it take the mindset of an actual Cyberattacker, and try to break down your walls of defenses ***to see where all gaps and weaknesses lie at, even those that are unknown.*** An illustration of it is below:


Pre-Engagement Interactions
Reconnaissance or OSINT
Threat Modeling & Vulnerability Identification
Exploitation
Post-Exploitation, Analysis, & Recommendations
Reporting

(SOURCE: 13).

The matric below reviews the major differences between the two methodologies:

Vulnerability Assessment	*Penetration Test*
Tests are passive	Tests are active
Tests are automated, no human intervention	Tests are primarily manual, lots of human intervention
Tests are short in time frame	Tests are much longer in time frame
Reports are provided to client, but not specific for remediative actions	Reports are provided to client, and are specific to remediative actions
Scans can be run on a continual cycle	Scanning is done only at point in time intervals due to their exhaustive nature
Tests are primarily done on digital assets	Tests are done on both physical and digital assets
Only known vulnerabilities are discovered	Both known and unknown vulnerabilities are discovered
Costs are affordable	Costs can be quite expensive
Only general tests are done	All kinds of tests are done, depending upon the requirements of the client

Traditionally, both Vulnerability Scanning and Penetration Testing have been done manually. Meaning, it is human beings that have manually carried out these kinds of tests, and reported back the results to the client in the form of a report. But given the advent of ChatGPT, many vendors who develop both of these tools are now implementing both AI and ML into them.

To a certain extent, this does have some benefits:

- Vulnerability Scanning, while automated for some time, can now be conducted with greater precision than ever before. In fact, it can even create a report with the findings and remediations automatically.

- When it comes to Penetration Testing, there are many steps and procedures that are involved. It is not simply a onetime deal like Vulnerability Scanning is. It can take days, and sometimes even weeks to conduct a thorough Penetration Test. The good news here is that AI and ML can be used to help automate some of the routine and mundane tasks, so that the Penetration Testing team can concentrate on the bigger picture at hand.

- Although the team has been trained to take the mind of a Cyberattacker when conducting a Penetration Test, there are still something things, such as targets, that may be overlooked. With AI and ML, the team will have more assurances that all possible attack scenarios have been planned beforehand. An example of this would be possibly using ChatGPT in these kinds of instances.

- When doing the actual Penetration Test and compiling the findings into a report for the client, the team will often rely upon their own hindsight, knowledge, and experience in order to come up with remediations. But by using AI and ML, the system can learn from previous attacks, and apply that repository to the new test that is being conducted. That way, the team will literally have a database of knowledge from which to glean from, which has been created on an automated basis.

- Asa, it has been described throughout this whitepaper, the Cyberattacker is getting more and more sophisticated in the way they launch their attacks. In fact, they are even starting to use AI and ML, but for nefarious purposes. The opposite can also be done, where the Penetration Testing team can use AI and ML to stay one step ahead of the proverbial cat and mouse game.

- Many vendors have now launched Penetration Testing platforms that are totally automated. One primary example of this is the "Node Zero", by Horizon3.ai. It is both automatic and autonomous, and the end user can conduct an unlimited amount of Penetration Tests for a flat fee, versus paying up to $30,000.00 for each Penetration Test.

But despite these advantages, the controversy still remains: Which is better, AI/ML based automation or human testing? Well honestly, it takes the best of both worlds in order to have an optimal Penetration Test to be done. In this regard, it is important to remember that AI and ML have their distinct limits, which from there, requires the involvement of human testers.

Proactive Monitoring & Remediation – Big Data & Generative AI

One offshoot of ChatGPT is a new field in AI and ML which is known as "Generative AI". A technical definition of it is as follows:

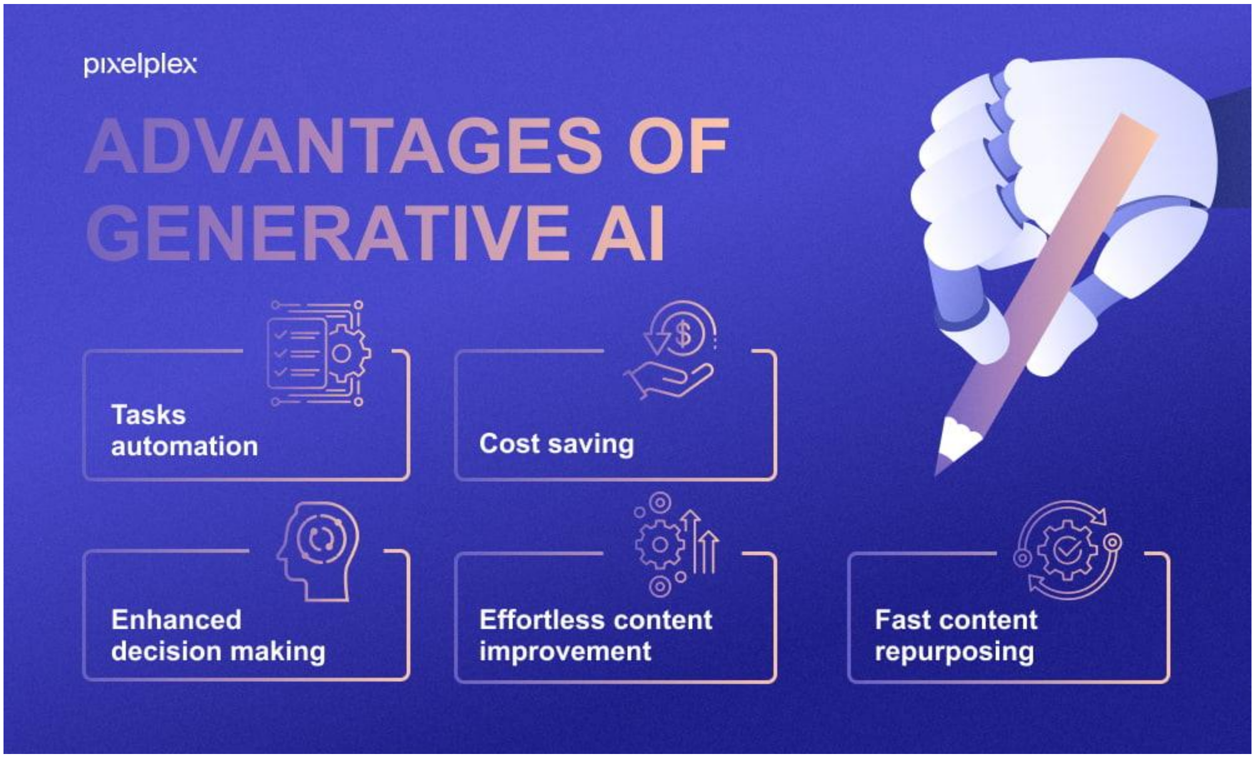
"Generative AI enables users to quickly generate new content based on a variety of inputs. Inputs and outputs to these models can include text, images, sounds, animation, 3D models, or other types of data."

(SOURCE: 14).

To explain this further in simpler terms, the traditional AI and ML models would only produce the outputs in terms of the data that were given to it. Over time, they would learn, and from there, try to extrapolate new projections just strictly based upon the queries that were posed to it. In other words, these models could really not make an attempt to "think outside of the box".

But Generative AI takes this one step further (or at least it attempts to) by trying to offer other types of solutions or outputs that was not asked of it in the original query. You may be asking at this point, "How is this possible"? Well, Generative AI tools like ChatGPT build a profile about you over time, and based on that, it tries to offer up other ideas or alternatives that you may not have asked of it to come up with.

An illustration Generative AI is provided below:



(SOURCE: 15).

Probably the best example of where Generative AI can be used like this is for content creation. All a writer literally has to do is ask ChatGPT to create content of a certain number of pages based upon the topic that it is given. And within minutes, he or she will have an eBook that can be published on Amazon.

These same techniques here as just described can also be applied to the world of Cybersecurity, when it comes to combing through all of the data that an IT Security team may receive. For example, depending upon the total number of network security devices that a business has, tons of data and information will be pouring in, on a daily basis. These are very often stored as log files. This data can consist of data about the total number of attempted and failed logins, and unusual activity that has exceeded the baseline, the total number of malicious data packets that have been intercepted, any spikes in network traffic in both the external and internal environments, etc.

Obviously, it would take a human being weeks or even months to parse through all of this data, and try to come up with any kind of reasonable conclusion from it. And in Cybersecurity, nobody has this kind of time to wait for. This is now where Generative AI can now come into play, and offer a huge helping hand. For instance, you can feed all of these log files into the Generative AI system, and literally within minutes, it can analyze the trends for you which you can view from one location.

From here, some of the most important things that you will want to take notice of are the spikes in network traffic, and the total number of attempted and rejected logins, and the times and dates that they have occurred. From here, you can then ask the system to create a profile for you so that you can infer your own observations from it.

For example, if the greatest number of failed logins happen after business hours, then this is the timeframe that your IT Security team will need to keep a close eye on for the short term. Also, depending upon the type of query that you pose to the Generative AI system, you can even ask it to find any correlations which are "hidden". In other words, these are trends which are not seen at first glance upon examination of the data, and could even take much longer if a human being had to parse through all of it.

But once again, within minutes, this system will search through all of the information that it has, and present these trends to you in an easy to decipher format. This has distinct advantages to it, and some of them are as follows:

- Given the discovery of these hidden trends, the data scientist will then be in a much better position to perhaps model future threat variants, and possibly even determine the trajectory path of existing threat variants.

- The information that is garnered from these hidden trends can be used to help to create and update existing Incident Response (IR) Plans, thus shortening the time it takes to respond to a security breach.

- These hidden trends can be used even for forensics purposes. For example, once a security breach has occurred, investigators are most interested in what is known as the "Latent Evidence". These are pieces of digital evidence (or even physical) that are hard to find the first time around. As a result of this, investigators will then keep combing the crime scene until all evidence has been collected. Once this has happened, the next major step is to piece together how the security exactly happened, and who the responsible party was. In this regard, these newly discovered, hidden trends can become the bridge to connect all of these dots together.

Now, if an organization is large enough, such as a Fortune 500 company, the amount of information and data that the IT Security will receive is gargantuan***. We are not talking about Gigabytes, we mean Petabytes***. There is a special term for this, and it is called "Big Data". A definition for it is as follows:

"Big data refers to data that is so large, fast or complex that it's difficult or impossible to process using traditional methods. The act of accessing and storing large amounts of information for analytics has been around for a long time. But the concept of big data gained momentum in the early 2000s when industry analyst Doug Laney articulated the now-mainstream definition of big data as the three V's:

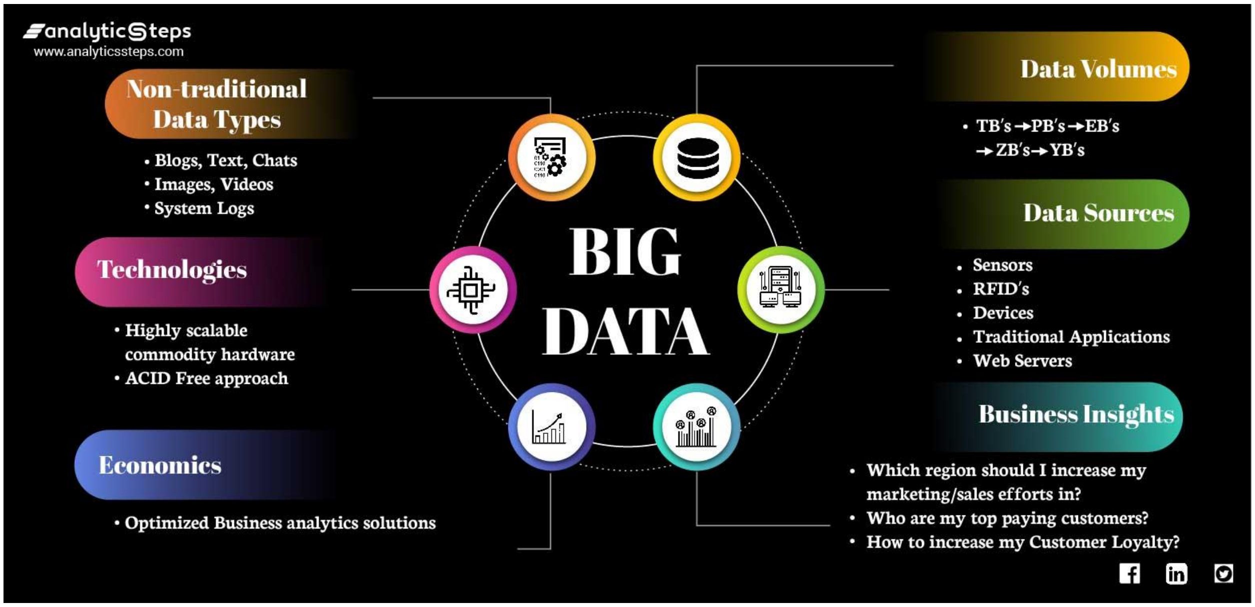
Volume. Organizations collect data from a variety of sources, including transactions, smart (IoT) devices, industrial equipment, videos, images, audio, social media and more. In the past, storing all that data would have been too costly – but cheaper storage using data lakes, Hadoop and the cloud have eased the burden.

Velocity. With the growth in the Internet of Things, data streams into businesses at an unprecedented speed and must be handled in a timely manner. RFID tags, sensors and smart meters are driving the need to deal with these torrents of data in near-real time.

Variety. Data comes in all types of formats – from structured, numeric data in traditional databases to unstructured text documents, emails, videos, audios, stock ticker data and financial transactions."

(SOURCE: 16).

An example of Big Data is below:


analyticSteps
www.analyticssteps.com
Non-traditional Data Types
• Blogs, Text, Chats
• Images, Videos
• System Logs
Technologies
• Highly scalable commodity hardware
• ACID Free approach
Economics
• Optimized Business analytics solutions
BIG DATA
Data Volumes
• TB's→PB's→EB's →ZB's→YB's
Data Sources
• Sensors
• RFID's
• Devices
• Traditional Applications
• Web Servers
Business Insights
• Which region should I increase my marketing/sales efforts in?
• Who are my top paying customers?
• How to increase my Customer Loyalty?

(SOURCE: 17).

A Fortune 500 company will most likely have offices on a global scale, and thus, multiple Generative AI systems will have to be created in order to process and analyze the Big Data on a daily basis. But by having this in place, an IT Security team will be able to get the answers that they need in a very rapid time frame.

Finally, with the traditional AI and ML models, the data scientist would have to create SQL commands in order to extract the results that were needed from the outputs. But Generative AI has changed all of this in the sense that all an end user has to do is merely ask or type in a query in plain, syntax format, and the answers will be provided almost immediately.

Proactive Monitoring & Remediation – Mitigating Against DDoS Attacks

It seems like there are many new threat variants coming out today, and that is actually true. The malicious payloads are far more destructive, covert, and stealthier than they ever used to be. But it is important to keep in mind here that these newer variants are simply "better mousetraps" from older versions. For example, look at Phishing. This is probably the oldest attack variant around, having evolved in the early 90s, with its first major impact on AOL.

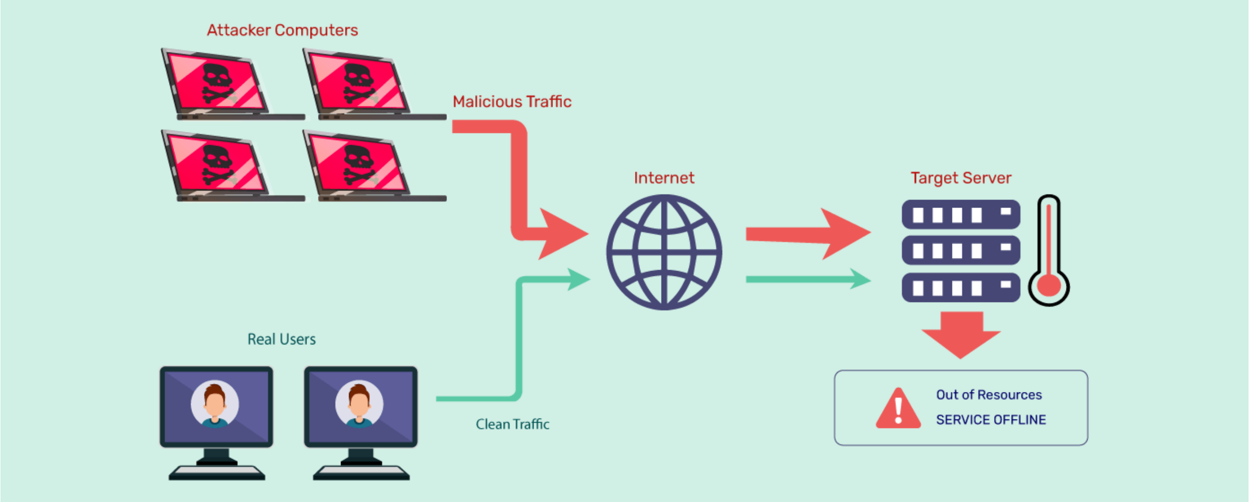
But, Cyberattackers have taken the fundamental core of what makes up Phishing, and have crafted it into a much nastier strain than ever before. An example of this is also what are known as "Distributed Denial of Service Attacks", or "DDoS" for short. A technical definition of it is as follows:

“A Denial-of-Service (DoS) attack is an attack meant to shut down a machine or network, making it inaccessible to its intended users. DoS attacks accomplish this by flooding the target with traffic, or sending it information that triggers a crash. In both instances, the DoS attack deprives legitimate users (i.e. employees, members, or account holders) of the service or resource they expected.”

(SOURCE: 18).

In simpler terms, this is where a Cyberattacker launches malformed data packets at a server, or even a group of servers, and keeps flooding them to the point where they cannot deliver services anymore to end users. The reason for this is that all of the processing and computer powers of the server are being used to process these data packets from which it is being hit on a nonstop basis.

An illustration of a DDoS attack is below:



(SOURCE: 20).

The unfortunate part of this is that DDoS attacks are once again ramping up; but the good news is that AI and ML can be used to give you early warning of a potential attack, thus mitigating your risk greatly in becoming a victim. Here are the different ways in which they both can be used:

1) <u>AI/ML can allow for real time monitoring</u>:

   While firewalls and routers are specifically designed to sniff out for malicious depending on packets, they won’t know what to look for until you put in the rules and permutations in place. This means that you have to keep up to par with the latest threats in network traffic on a daily basis, which of course nobody has the time to do this. But by using AI/ML into your overall network security posture, these tools can keep an eye on a 24 X 7 X 365 basis on your network traffic, and alert you on a real time basis of an imminent DDoS threat. Another benefit here is that AI/ML can even update the rules and permutations in your network security devices with what it is learning, so you will not have to do this on a manual basis.

2) <u>AI/ML and be used for both Predictive Analytics and Behavioral Analysis</u>:

   When you first set up your network security perimeter, one of your main tasks is to come up with a baseline profile of what is considered to be “normal traffic”. This can be time consuming task, and it will have to be repeated over and over again on a regular schedule. But given the powerful algorithms that AI and ML both have now, they can learn what is a normal traffic pattern, and from there set up the appropriate baselines. Then, it can start to detect any events or activities that fall outside of this. Remember, the Cyberattacker of today does not launch a DDoS attack just out of the blue. They take their time to study their targets, and from there, determine what the weak spots are in the servers. AI and ML can pick up on this, and will alert you to this fact.

3) AI/ML can be used for Correlation Analysis:

   It would be a nearly impossible task for a human being to analyze each and every data packet that comes through, and try correlate the information and data that are found in them to other data packets. But AI and ML for the most part can handle this kind of task, but of course the algorithms will have to be created in such a way that it can accomplish this task. For example, they can ferret out the good and bad data packets, and segregate both categories. From here, they can learn and create what constitutes acceptable and malformed data packets. This can then be used to examine future data packets, and determine if there are any relationships which exist. To take even one step further, what is learned here from this Correlation Analysis can even be used to project what future DDoS attacks could be like into the future, and even provide clues into their attack signatures.

4) AI/ML can be used for Automated Responses:

   It is important to note that AI and ML can be used so much more than for data packet analysis. They can be used to automate your desired response to a DDoS attack. For example, if your systems pick up that an attack is imminent, the AI and ML tools can then be used to shut down the required systems for a short period of time until the threat fades off. It can even send out notifications in real time to all employees that certain resources will be offline for a short duration because of an imminent security breach. Another advantage here is that based upon the analytics that the AI and ML tools give you, it is quite possible that you can isolate these malformed data packets into a sandboxed environment, and study the malicious payload. From here, you can then feed into the AI and ML tools what you have discovered, in an effort to enrich their own databases for future Correlation Analyses.

Proactive Monitoring & Remediation – Effective Asset Discovery

As a business owner, or even as a CISO, one of the first projects that you will want to endeavor through is that of conducting a Risk Assessment Study. But before you can do this, you first need to know what all of your assets are in both your IT and Network Infrastructures. This is known as “Asset Discovery”. It can be specifically defined as follow:

“The process of locating, identifying, and indexing the totality of an organization’s IT assets. This means both hardware and software.”

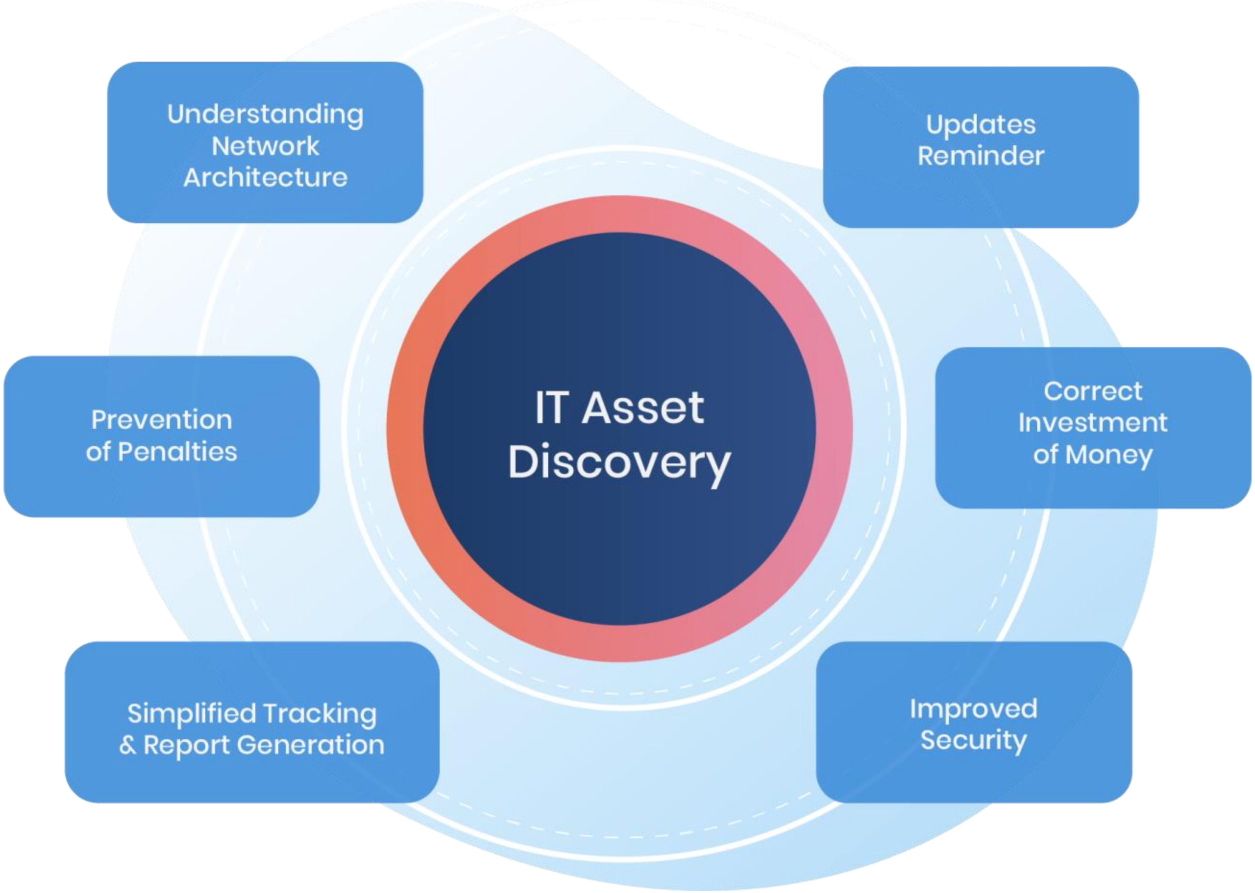
(SOURCE: 20).

Examples of IT related assets include the following:

- “Laptops
- Tablets
- PCs
- Smartphones
- Servers
- Internet of Things (IoT) devices like printers
- Switches
- Routers
- In terms of software, an organization’s network may have:
- Software libraries (.dll, .ocx, etc.)
- Unauthorized software
- Authorized software
- Executable (.exe) files and software scripts (.py, .ps1, etc.)
- Examples of mission-critical cloud resources could be:
- Code repositories
- Storage buckets
- Databases-as-a-server (DBaaS)
- Containers
- Virtualized network devices
- VMware, Hyper-V, EC2 instances, and other virtual servers

(SOURCE: 20).

An illustration of effective Asset Discovery is below:


IT Asset Discovery
Understanding Network Architecture
Updates Reminder
Prevention of Penalties
Correct Investment of Money
Simplified Tracking & Report Generation
Improved Security

(SOURCE: 21).

Now imagine if you and your IT Security had to go through and map find all of these assets using a manual approach. This of course would be a nightmare, and take forever to complete. The good news here is that AI and ML can be used to help accomplish this process in just a very short time period. The way it does this is by scanning everything in your IT environment, and from there, categorizing and labeling all of the assets it has discovered.

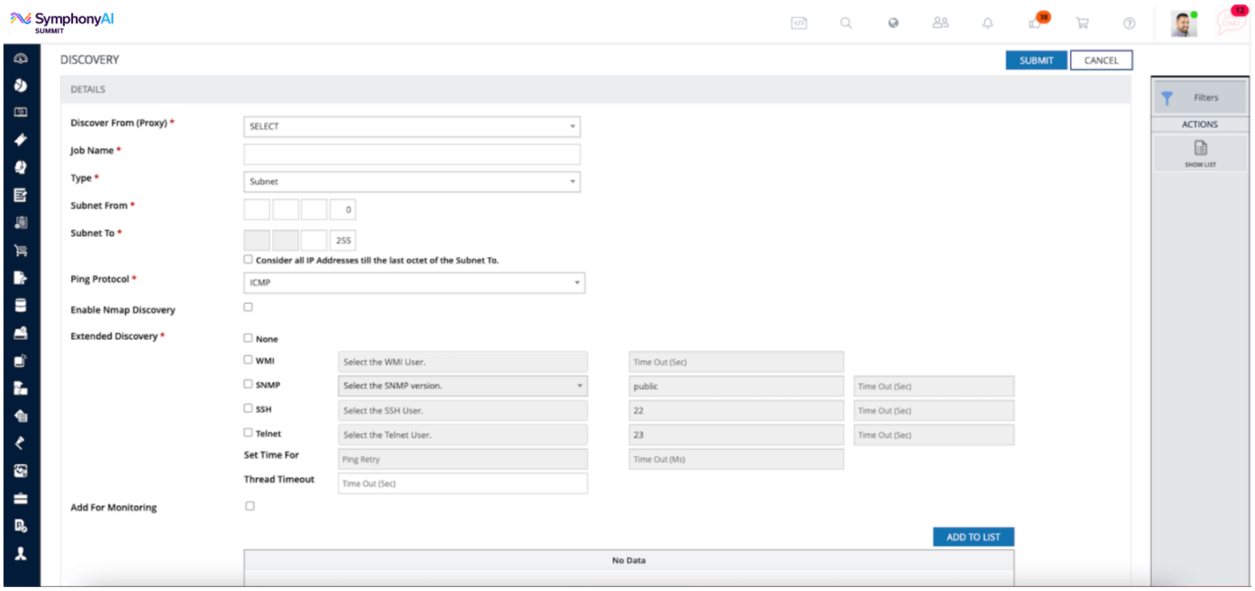
Some of the other advantages of using AI and ML in this fashion include the following:

1) <u>Updates can happen on a real time basis</u>:

   Whether you still have a brick-and-mortar presence or even if it is virtual based with a remote workforce, you will always be adding new equipment to your inventory list. By using AI and ML you can update this list on a real time basis, thus avoiding the need to waste time and doing it manually. But in this regard, it is also very important to remember that IT assets are not just physical assets. They can also be virtual, and exist in a Cloud based platform, such as Microsoft Azure. These also need to be included in any asset discovery process.

2) <u>Everything can be seen from a central place</u>:

   By using AI and ML in lieu of Excel spreadsheets, you will be able to see all of your digital assets from one single console. This will allow for better management of all of the devices that are being used, and you can even examine how one asset is being used in correlation with other assets located throughout your business, across the different departments that you may have. An example of this is illustrated below:



DISCOVERY

DETAILS

Discover From (Proxy) * SELECT

Job Name *

Type * Subnet

Subnet From * □ □ □ 0

Subnet To * □ □ □ 255

□ Consider all IP Addresses till the last octet of the Subnet To.

Ping Protocol * ICMP

Enable Nmap Discovery □

Extended Discovery *

□ None			
□ WMI	Select the WMI User.	Time Out (Sec)	
□ SNMP	Select the SNMP version.	public	Time Out (Sec)
□ SSH	Select the SSH User.	22	Time Out (Sec)
□ Telnet	Select the Telnet User.	23	Time Out (Sec)
Set Time For	Ping Retry	Time Out (Ms)	
Thread Timeout	Time Out (Sec)		

Add For Monitoring □

(SOURCE: 22).

3) Realize improved metrics:

One of the most important metrics in Cybersecurity is what is known as the "Mean Time To Resolution", also known as "MTTR". It can be defined as:

"MTTR (meaning time to resolve) is the average time it takes to fully resolve a failure. This includes not only the time spent detecting the failure, diagnosing the problem, and repairing the issue, but also the time spent ensuring that the failure won't happen again."

(SOURCE: 23).

By having a centralized dashboard that is powered by both AI and ML, you will be able to see very quickly which of your assets is not functioning up to its optimal level. From here, you can then take the appropriate, remediative steps which are needed. The result here is that you will have very little downtime (if any) in terms of productivity.

4) Maintenance costs are lowered:

Most businesses don't' realize its time to either maintain or completely replace a device until their employee tells them or the IT Department brings it up. When this happens, the cost of finding a suitable replacement can quickly escalate, especially if it is deemed to be "out of life". So why take this chance when you can create a centralized dashboard that can give you these necessary updates on a real time basis, thus allowing you to plan on a proactive basis?

5) Better compliance:

By seeing what is happening to your digital assets from a bird's eye view, you can see which ones are more vulnerable to a security breach. From here, you can then wither upgrade or implement newer controls so that you stay with the data privacy laws of the GDPR, CCPA, HIPAA, etc. But this will be based upon completing the Risk Assessment Study, as eluded before. AI and ML can be used here in this instance also, and will be the focal point for a future whitepaper.

Conclusions

Overall, this whitepaper has provided an overview into what both AI and ML are all about, and how they both can be used to help improve your security posture. But remember, they will never replace humans. At best, they will inly augment existing processes by helping to automate them. Also, never try to rely on AI and ML alone to solve your Cyber needs. It takes both a combination of technology and people, and finding that balance is what is of paramount importance.

Finally, if you have any questions about the concepts presented in this whitepaper, or are interested in learning more about how you can deploy AI and ML tools in your business, contact us today.

Sources

1) https://www.sas.com/en_us/insights/analytics/what-is-artificial-intelligence.html
2) https://www.datamation.com/artificial-intelligence/what-is-artificial-intelligence.html
3) https://cybersecurityventures.com/jobs/
4) https://www.forbes.com/sites/cognitiveworld/2019/06/19/7-types-of-artificial-intelligence/#301ec3a6233e
5) https://www.expertsystem.com/machine-learning-definition/
6) https://searchenterpriseai.techtarget.com/definition/machine-learning-ML
7) https://www.proofpoint.com/us/threat-reference/alert-fatigue
8) https://www.ibm.com/topics/siem
9) https://www.manageengine.com/products/eventlog/manageengine-siem-whitepaper.html
10) https://www.techtarget.com/searchsecurity/definition/vulnerability-scanning
11) https://www.prplbx.com/resources/blog/vulnerability-scanning/
12) https://www.cisco.com/c/en/us/products/security/what-is-pen-testing.html
13) https://cipher.com/blog/a-complete-guide-to-the-phases-of-penetration-testing/
14) https://www.nvidia.com/en-us/glossary/data-science/generative-ai/
15) https://pixelplex.io/blog/generative-ai/
16) https://www.sas.com/en_us/insights/big-data/what-is-big-data.html
17) https://www.analyticssteps.com/blogs/top-10-big-data-technologies-2020
18) https://www.paloaltonetworks.com/cyberpedia/what-is-a-denial-of-service-attack-dos
19) https://www.indusface.com/learning/what-is-a-ddos-attack/
20) The process of locating, identifying, and indexing the totality of an organization's IT assets. This means both hardware and software.
21) https://blog.invgate.com/best-practices-it-asset-discovery-and-inventory-management
22) https://www.symphonysummit.com/products/it-asset-management/it-asset-discovery-2/?utm_source=google&utm_medium=cpc&utm_campaign=Asset-Management-

US&utm_term=asset%20discovery%20tools&gad=1&gclid=EAIaIQobChMIj7vqmK_7gAMVyyStBh1p1gRMEAAYASAAEgKunPD_BwE

23) https://www.atlassian.com/incident-management/kpis/common-metrics#:~:text=MTTR%20(mean%20time%20to%20resolve,failure%20won't%20happen%20again.