Image courtesy: Asset Guardian (https://www.assetguardian.com/how-ai-is-changing-the-cyber-security-landscape/)

AI in Cybersecurity: The Good, the Bad, and the In-between.

Research Interest

Intersection of

- Cybersecurity
- Digital Forensics
- Data Science
- Machine Learning

Current Research

Focuses on extracting and analyzing data patterns created by deleted and decaying digital files, and their application in digital forensics investigations to reconstruct previous user activity.

O. Cheche Agada, Ph.D.

Assistant Professor
Information Sciences and Technology
College of Engineering and Computing
George Mason University

- Ph.D. in Information Technology from George Mason University.
- MSc in Digital Forensics from George Mason University.
- MSc in Computer and Information Sciences from Southern Arkansas University.
- BSc in Computer Science from the University of Lagos, Nigeria.

Agenda

What is Cybersecurity?

Current Trends in Cybersecurity

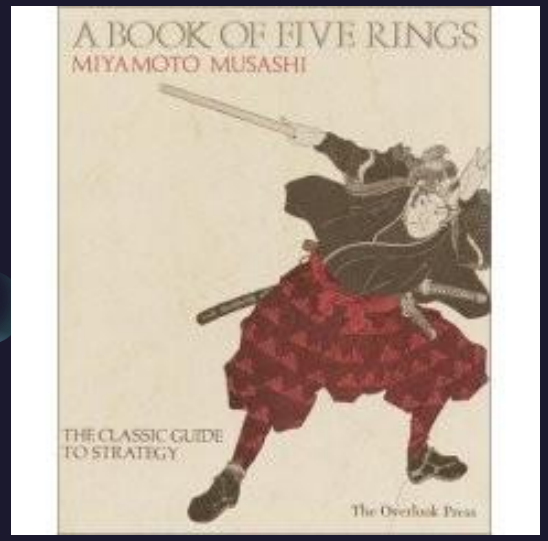
What is Artificial Intelligence?

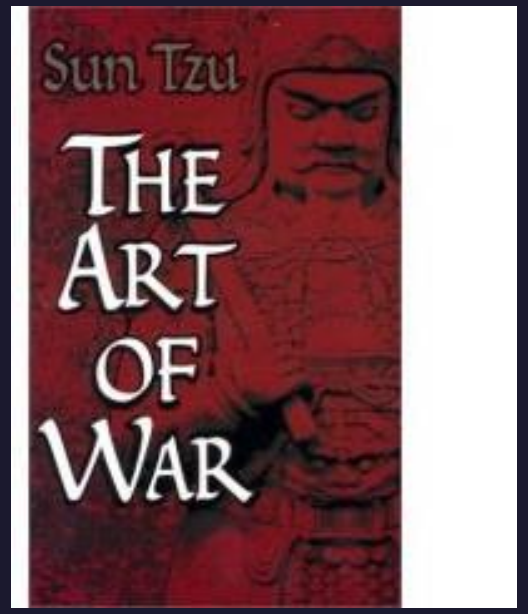
AI in Cybersecurity

- The Bad
- The Good
- The In-Between
- The Future

What is Cybersecurity?

What is Cybersecurity?





"The art of war teaches us to rely not on the likelihood of the enemy's not coming, but on our own readiness to receive him; not on the chance of his not attacking, but rather on the fact that we have made our position unassailable."

Sun Tzu, The Art of War

"Do not figure on opponents not attacking; worry about your own lack of preparation."

Miyamoto Mushashi, A Book of Five Rings

What is Cybersecurity?

CISCO

- Cybersecurity is the practice of protecting systems, networks, and programs from digital attacks. These cyberattacks are usually aimed at accessing, changing, or destroying sensitive information; extorting money from users via ransomware; or interrupting normal business processes.

WIKIPEDIA

- Computer security, cybersecurity, digital security or information technology security (IT security) is the protection of computer systems and networks from attacks by malicious actors that may result in unauthorized information disclosure, theft of, or damage to hardware, software, or data, as well as from the disruption or misdirection of the services they provide.

CISA

- Cybersecurity is the art of protecting networks, devices, and data from unauthorized access or criminal use and the practice of ensuring confidentiality, integrity, and availability of information.

What is Cybersecurity?

Cybersecurity vs Information Security

https://youtu.be/CKaI57Upbpg

What is Cybersecurity?

Basic Concepts

CONFIDENTIALITY

- Assurance that information is not disclosed to unauthorized persons, processes, or devices. Confidentiality covers data in storage, during processing, and in transit. Sources: Confidentiality measures the attacker's ability to obtain unauthorized access to information from an application or system (NIST)

INTEGRITY

- The term 'integrity' means guarding against improper information modification or destruction, and includes ensuring information non-repudiation and authenticity. A property whereby data has not been altered in an unauthorized manner since it was created, transmitted, or stored (NIST).

AVAILABILITY

- The timely, reliable access to data and information services for authorized users. The state that exists when data can be accessed or a requested service provided within an acceptable period of time. It measures an attacker's ability to disrupt or prevent access to services or data (NIST).

What is Cybersecurity?

Basic Concepts

AUTHENTICATION

- Verifying the identity of a user, process, or device, often as a prerequisite to allowing access to resources in an information system. The process of proving the claimed identity of an individual user, machine, software component or any other entity. (NIST)

NON-REPUDIATION

- The inability to deny responsibility for performing a specific act. Assurance that the sender of information is provided with proof of delivery and the recipient is provided with proof of the sender's identity, so neither can later deny having processed the information. (NIST).

ACCESS CONTROL

- The set of procedures and/or processes that only allow access to information in accordance with pre-established policies and rules. To ensure that an entity can only access protected resources if they have the appropriate permissions based on the predefined access control policies. (NIST).

What is Cybersecurity?

Basic Concepts

Threat:

A threat represents the potential for security to be compromised. Anything that represents a possible compromise of security is considered a threat. It does not require a clear action, although it may include one.

- Natural events – Caused by nature
- Human errors – Caused by people but unintentionally
- Attacks – Caused by people intentionally. An attack requires intent, which requires a person/persons.

What is Cybersecurity?

Basic Concepts

Attack:

An attack may or may not be successful.

A successful attack (or some other type of threat action) may be called

- an incident
- a compromise
- a breach
- an exploit

An unsuccessful attack may provide useful information to both attackers and defenders. (Frequently used in SQL injection probing.)

What is Cybersecurity?

Basic Concepts

Question

Why do we have security problems?
Why can't we have perfectly secured systems?
Why do we have security breaches?

What is Cybersecurity?

Basic Concepts

Security is not absolute.

There is always a trade-off between security and usability:

A perfectly secure system would be unusable.
A perfectly usable system would have no security.

"Those who would give up essential Liberty, to purchase a little temporary Safety, deserve neither Liberty nor Safety."
Benjamin Franklin

"Information security is a "well-informed sense of assurance that the information risks and controls are in balance".
James Anderson

What is Cybersecurity?

Basic Concepts

The Problem with security.

There is a misunderstanding of the security problem.

Security is not a technology problem, it is a business problem.

We often hear how important security is, but we don't always understand why. Security is important because it helps to ensure that an organization is able to continue to exist and operate in spite of any attempts to steal its data or compromise its physical or logical elements. Security should be viewed as an element of business management rather than an IT concern. In fact, IT and security are different. Information technology (IT) or even information systems (IS) is the hardware and software that support the operations or functions of a business. Security is the business management tool that ensures the reliable and protected operation of IT/IS. Security exists to support the objectives, mission, and goals of the organization. – Mike Chapple/James Michael Stewart/Darril Gibson

What is Cybersecurity?

Skillset and Careers

The NICE Framework--NIST Special Publication 800-181--categorizes and describes cybersecurity work through a taxonomy and common lexicon.

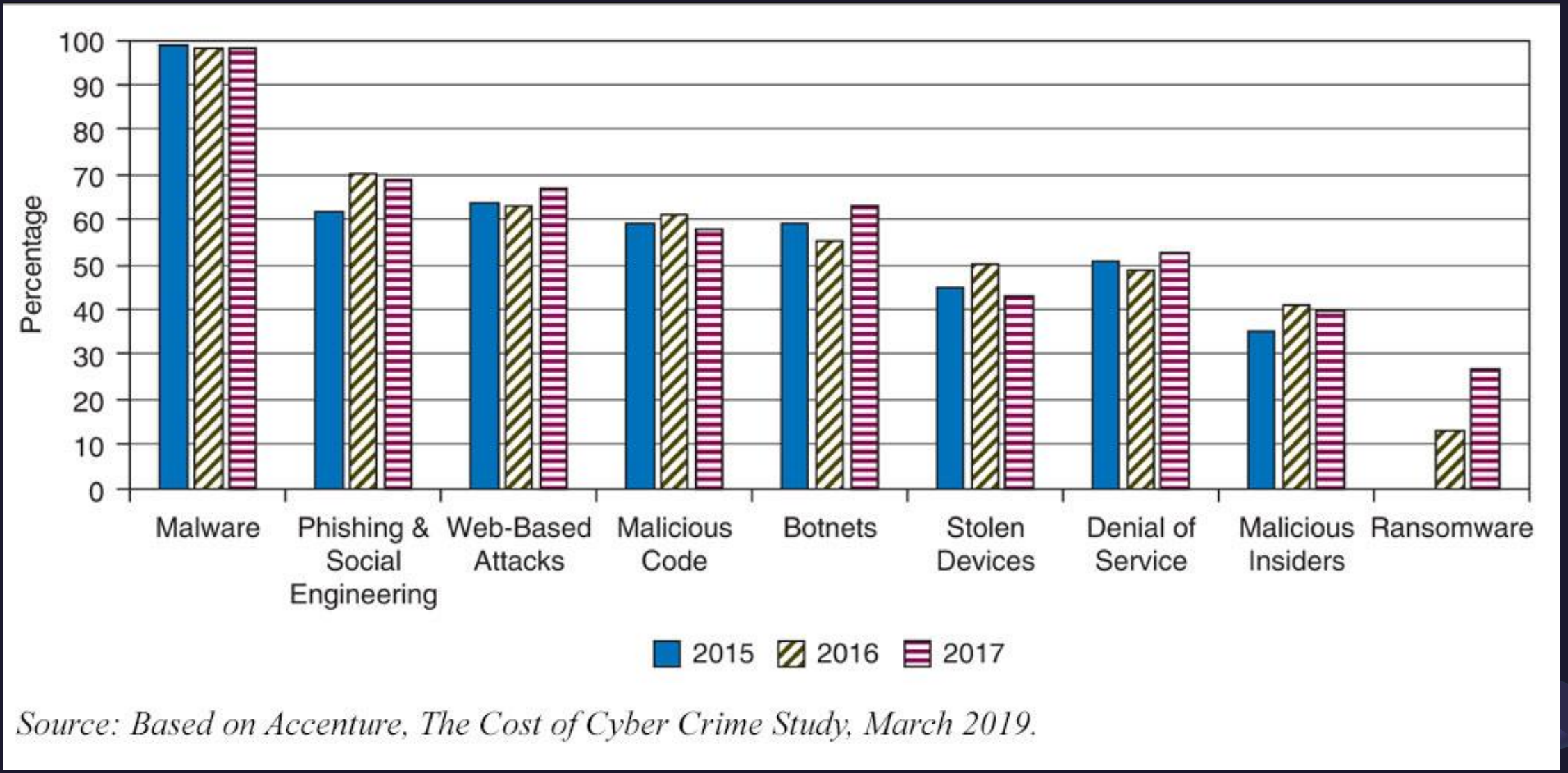
The NICE Framework is comprised of the following components:

- Categories (7) – A high-level grouping of common cybersecurity functions.
- Specialty Areas (33) – Distinct areas of cybersecurity work.
- Work Roles (52) – The most detailed groupings cybersecurity work comprised of specific knowledge, skills, and abilities required to perform tasks in a work role.

https://www.nist.gov/itl/applied-cybersecurity/nice/resources/nice-cybersecurity-workforce-framework

Current Trends in Cybersecurity

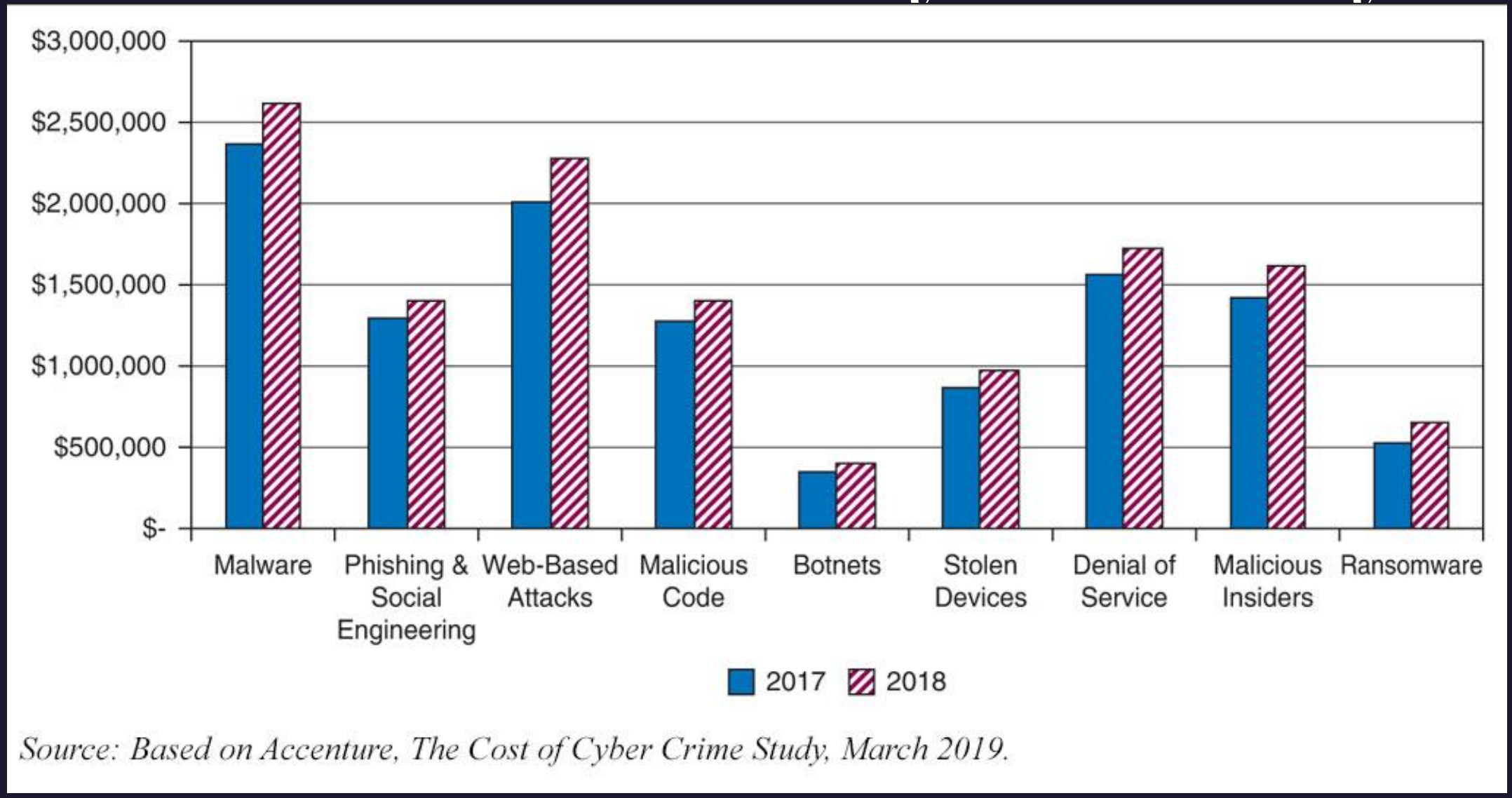
Current Trends in Cybersecurity



Source: Based on Accenture, The Cost of Cyber Crime Study, March 2019.

Percentage of Companies Experiencing Attacks by Attack Type

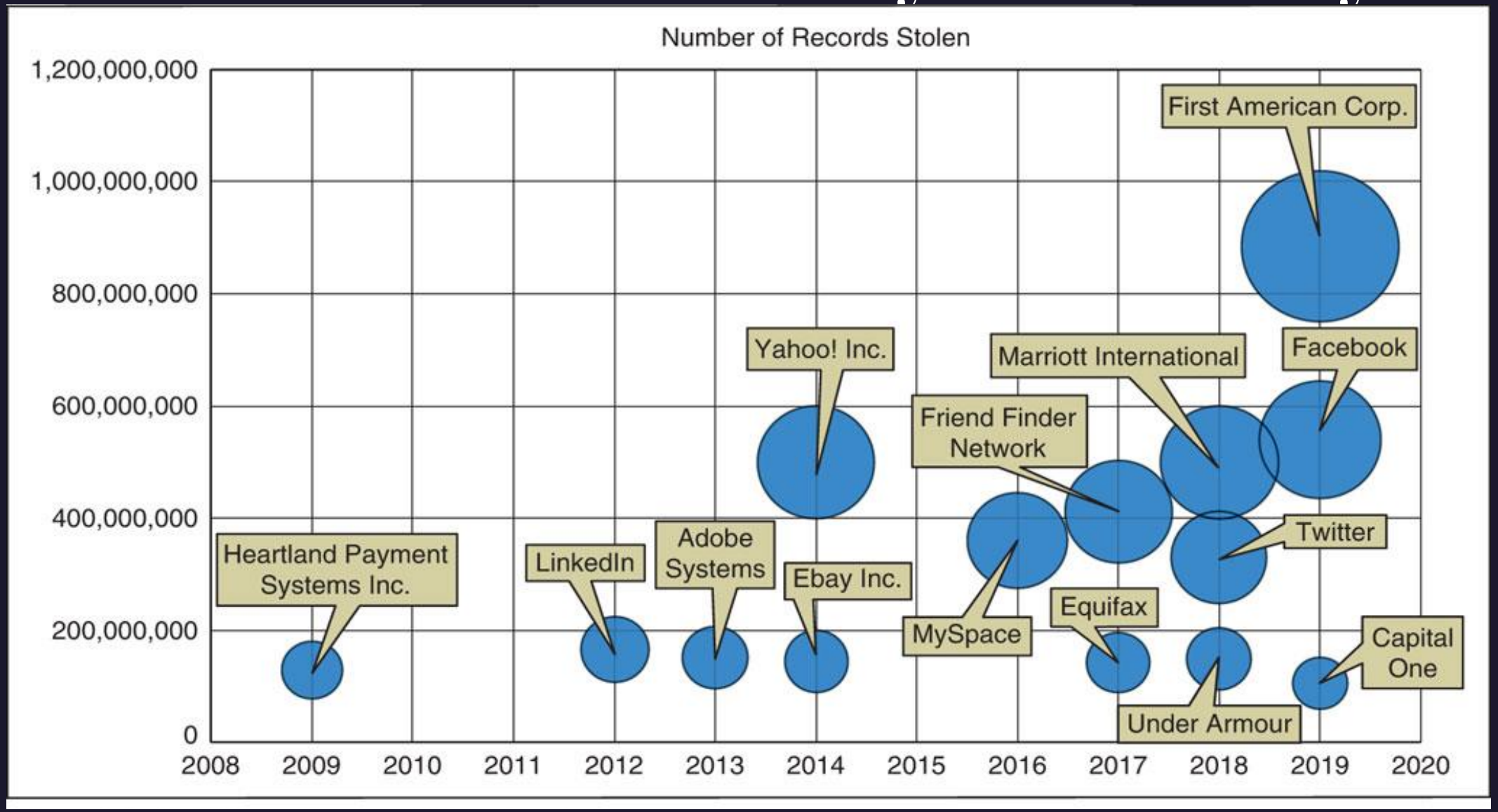
Current Trends in Cybersecurity



Source: Based on Accenture, The Cost of Cyber Crime Study, March 2019.

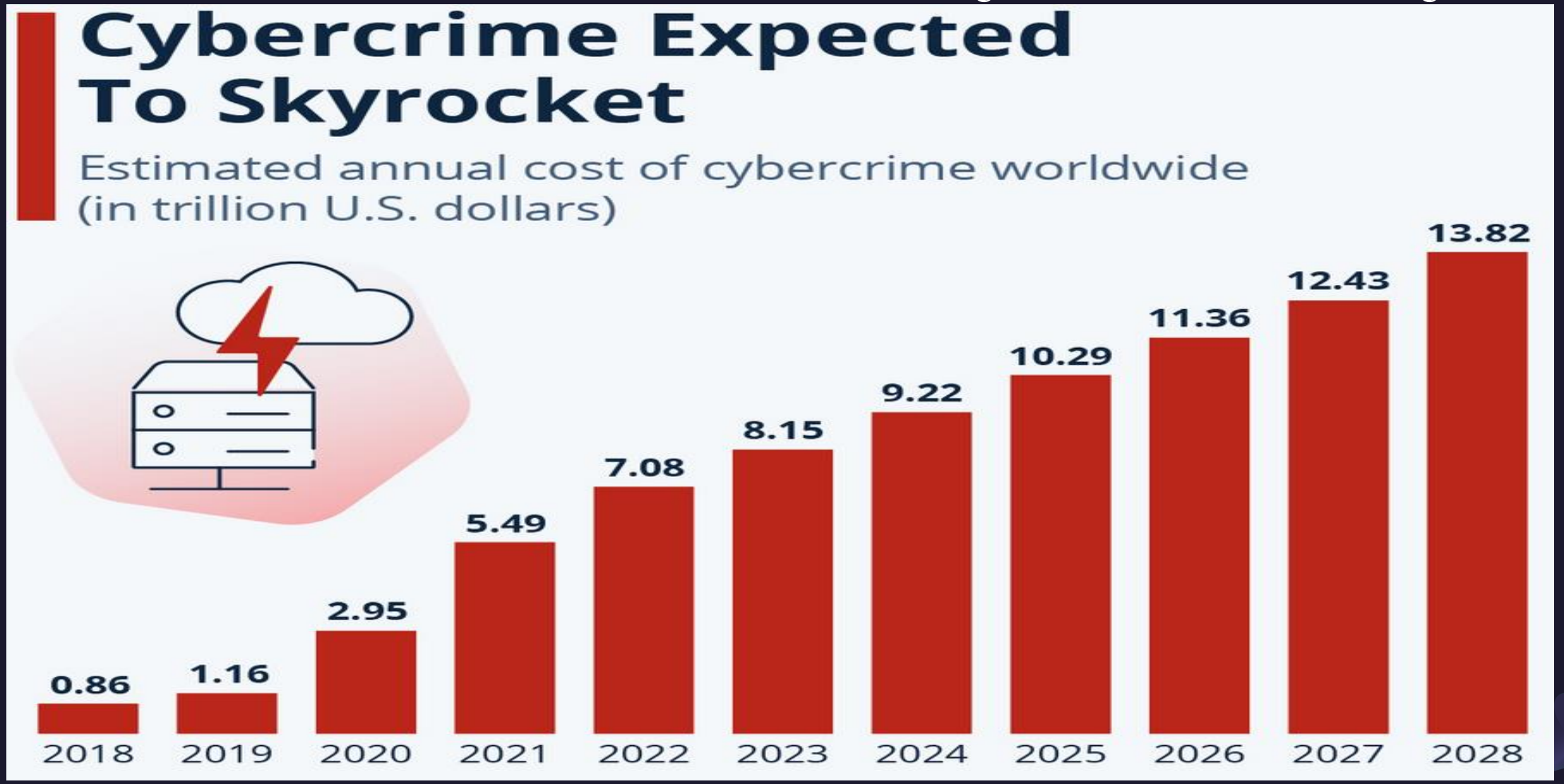
Average Annual Computer Crime Costs by Attack Type

Current Trends in Cybersecurity



Data Breaches

Current Trends in Cybersecurity



https://www.statista.com/chart/28878/expected-cost-of-cybercrime-until-2027/

Estimated Annual Cost of Cybercrime

Current Trends in Cybersecurity

Threat Actors

Here's a quick review of cyber threat actors we care about:

Financially motivated:

- Cybercriminals – Their skills range across a broad spectrum, but by-and-large they are fairly capable using advanced capabilities.
- State-sponsored actors?

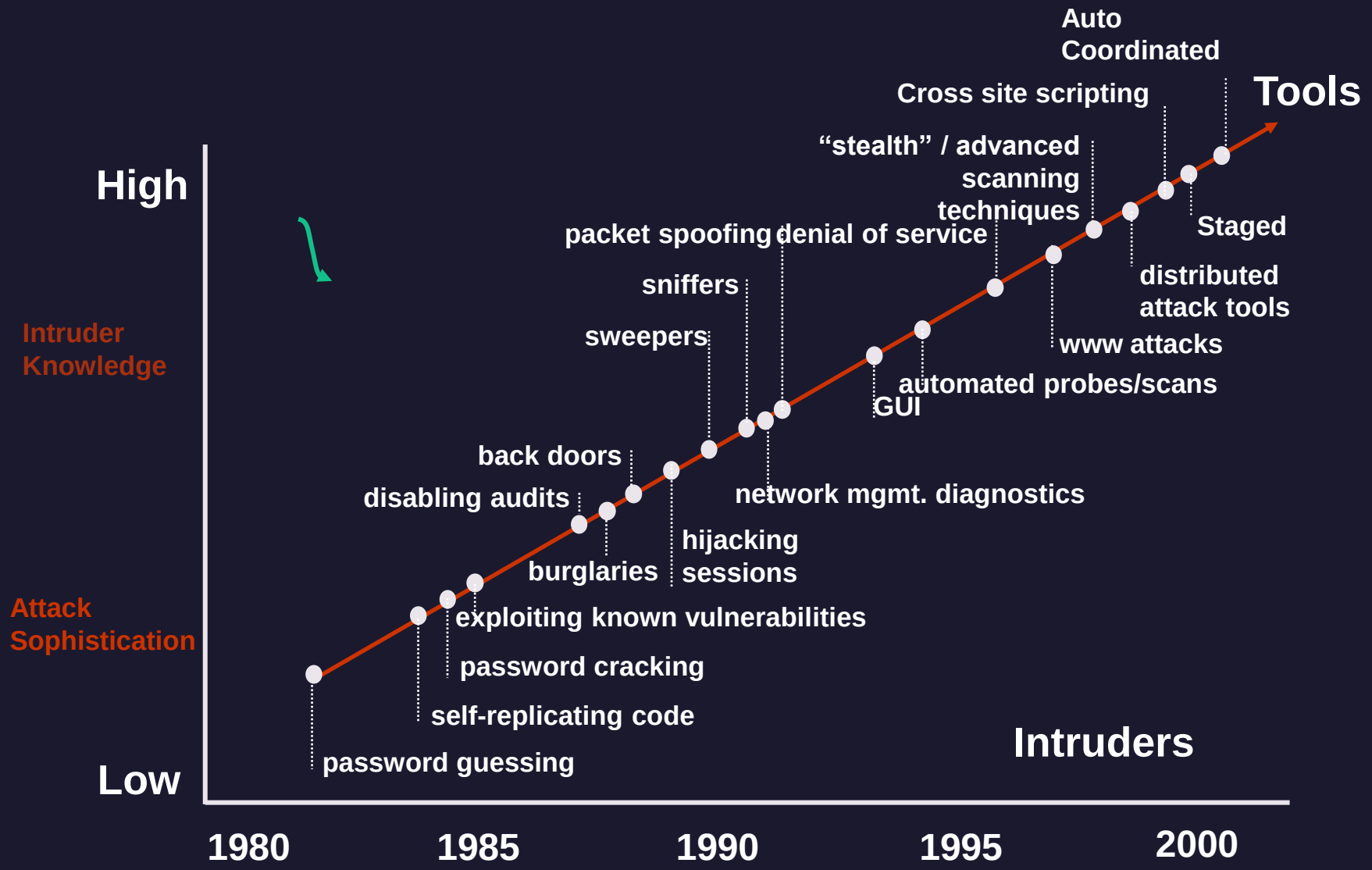
Ideologically motivated:

- Hacktivists – typically rely upon DDoS attacks to take down web services related to a certain event or group. Anonymous is one of the most well-known hacktivist collectives.
- Terrorists

Intelligence/Intellectual Property:

- Competitors
- Insiders
- State-sponsored aka "Advanced Persistent Threats" (APTs)

Current Trends in Cybersecurity



Attack Sophistication vs. Intruder Technical Knowledge

Current Trends in Cybersecurity

Cyber Attack Trends

- Attacks are happening more frequently and attacking a larger swath of devices.
- Automated tools found online allow individuals with minimal computer skills the ability to cause serious damage to targets.
- Rather than random attacks, attacks are targeting high-gain sources.
- Since 2013, there has been an emergence of Advanced Persistent Threats (APT) where hackers gain unauthorized access and remain undetected for a long period of time.

Current Trends in Cybersecurity

Cyber Attack Trends

High points from Symantec's June 2017 "Living off the Land and Fileless Attack Techniques" report:

1) Attackers are increasingly making use of **tools already installed on targeted computers or are running simple scripts and shellcode directly in memory**. Creating less new files on the hard disk means less chance of being detected by traditional security tools and therefore minimizes the risk of an attack being blocked.

2) Malicious scripts are hidden inside the registry or Windows Management Instrumentation (WMI) in order to achieve a **stealthy fileless persistence method** on a compromised computer. System and dual-use tools are frequently used in order to gather information about a freshly compromised system i.e. PowerShell. These tools have also been used during lateral movement or to exfiltrate stolen data. This activity blends in with normal system administration work.

3) Attackers are reverting back to these simple but proven methods, as it is getting more cost intensive to find reliably exploitable vulnerabilities. **Often a spear-phishing attack with some social engineering can be just as successful at achieving the attackers' goals**.

Current Trends in Cybersecurity

Recent Cyber Attacks Worth Mentioning

Software Supply Chain Attack against security tool CCleaner (Sept 2017)

Destructive Global Cyber Attacks: NotPetya (June 2017)

- **Infected systems in over 65 countries disproportionately affecting those in system found in Ukraine**. The malware masqueraded as ransomware, but really just destroyed key system files rendering the systems inoperable.

Global Ransomware Cyber Attacks – WannaCry (May 2017)

- **Spread to over 200,000 computers worldwide in a matter of hours, extorting victims for ~$300 to decrypt files**. Over 60 UK hospitals were affected. Note: Microsoft released a patch to protect users against the vulnerability exploited by WannaCry a month before the attack

Current Trends in Cybersecurity

Recent Cyber Attacks Worth Mentioning

Global Denial-of-Service Attack – DynDNS/OVH/Krebs DDoS attacks (Mirai Botnet) (~Sept – Oct 2016)

Hacks that Influence National Elections - Democratic National Committee Hack (July 2016)

- Alleged Russian hackers compromised the computer network of the Democratic National Committee, releasing over 20,000 emails in a smear campaign against Hilary Clinton.

Targeted Attack Against Critical Infrastructure – Black Energy (December 2015)

- **Multiprong attack against Ukraine's electric sector, which left areas of the country with out power for several days. The attack impacted 1 company in 3 locations**. The attackers took control of the mouse on the HMI and DDoSed the help desk to delay technicians from determining an issue was occurring.

Current Trends in Cybersecurity

Recent Cyber Attacks Worth Mentioning

Famous Ransomware Attack

- Fulton county, Georgia – Hackers threaten to release Trump documents from Georgia case if they don't get a ransom by Thursday https://www.businessinsider.com/trump-georgia-documents-ransom-threat-fulton-county-hack-lockbit-2024-2

Current Trends in Cybersecurity

Prelude on Offensive Cyber Terminology

There is a cause-and-effect relationship between a vulnerability and an exploit. i.e. an attacker exploits a vulnerability to cause some action.

Zero-day Exploit

- A zero-day exploit takes advantage of a previously unknown vulnerability that only becomes known once an attacker has been caught using it.

N-day Exploit aka Common Vulnerability and Exposure (CVE)

- A n-day exploit, often referenced by its CVE number, is an exploit that takes advantage of a publicly known vulnerability.

https://cve.mitre.org/

Current Trends in Cybersecurity

Cyber Security Industry Trends

- (~2017) "Hunt Teams" is an emerging trend. The term was first coined in ~late 2015 by Robert Lee (SANS) with company Sqrrl leading field.
- (~2017) Detection technology in security products is improving, harnessing advanced computing power and machine learning models. For instance, CrowdStrike's Falcon View and FireEye's products are top-notch.
- (~2015-2017) The cyber security industry is consolidating with several niche vendors beginning to merge the offer more comprehensive solutions to prospective clients. Typically, it is an incident response company merging with an end-point detection company merging with a forensics company that then sells threat intelligence that offers end point protection and can do on-site remediation post-breach.
- (~2015-2017)Operating Systems are getting harder to exploit…if fully patched!
- (~2015-2016) Security automation/DevOps
- (~2014-2015) Thanks to Splunk, security and event management (SIEM) products are aggregating log files and threat feeds into one place and have created a robust way to query for threats occurring within a network.
- (~2016) Thanks again to Splunk, which has included visualization into its product to help security professionals illustrate to their management the threat its network faces.
- (~2013+) Pivoting off of open source data to find additional indicators of compromise like the VirusTotal data set.

Current Trends in Cybersecurity

Other Cyber Trends Worth Mentioning

- There is a trend occurring where the United States through its Department of Justice has been issuing arrest warrants for suspected state-sponsored hackers through public indictments:

- 12/20/2018 - Two Chinese Hackers Associated With the Ministry of State Security Charged with Global Computer Intrusion Campaigns Targeting Intellectual Property and Confidential Business Information

- 11/28/2018 - Two Iranian Men Indicted for Deploying Ransomware to Extort Hospitals, Municipalities, and Public Institutions, Causing Over $30 Million in Losses

- 10/30/2018 - Chinese Intelligence Officers and Their Recruited Hackers and Insiders Conspired to Steal Sensitive Commercial Aviation and Technological Data for Years

- 10/4/2018 - U.S. Charges Russian GRU Officers with International Hacking and Related Influence and Disinformation Operations

- 9/6/2018 - North Korean Regime-Backed Programmer Charged With Conspiracy to Conduct Multiple Cyber Attacks and Intrusions

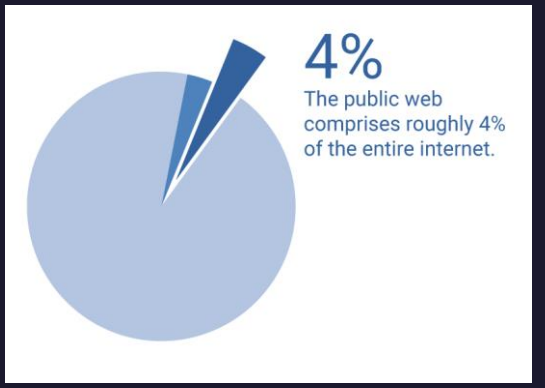
- 3/23/2018 - Nine Iranians Charged With Conducting Massive Cyber Theft Campaign on Behalf of the Islamic Revolutionary Guard Corps

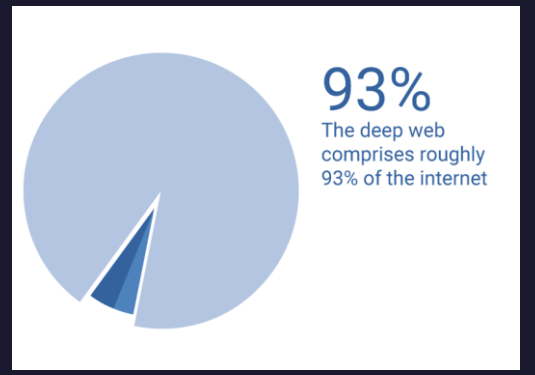
Current Trends in Cybersecurity

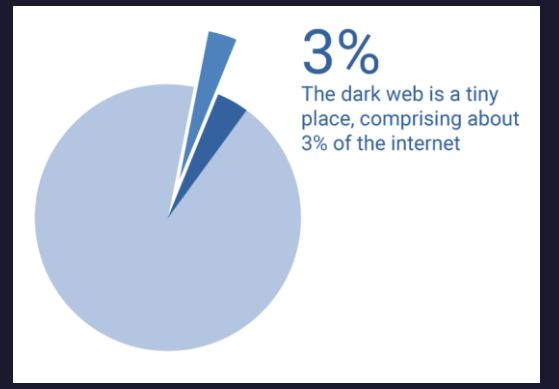
Dark Web/Dark Net Activities

"If the surface web is the tip of the iceberg and the deep web is what's below the water, then the darknet is what you'll find deep in the blackest waters below. The darknet is the network itself, whereas the dark web is the content that is served up on these networks."

https://www.cnet.com/news/darknet-dark-web-101-your-guide-to-the-badlands-of-the-internet-tor-bitcoin/







http://www.experian.com/blogs/ask-experian/what-is-the-dark-web/

Current Trends in Cybersecurity

Dark Web/Dark Net Activities

https://www.ice.gov/features/darknet

Current Trends in Cybersecurity

A Vishing Call

? Questions ?

Image courtesy: ICT Mentors Solutions Limited (https://www.linkedin.com/pulse/ai-cyber-security-ict-mentors-solutions-exudf/?trk=articles_directory)

What is Artificial Intelligence?

What is Artificial Intelligence?

AI Defined

- Artificial Intelligence is a branch of computer science that endeavors to replicate or simulate human intelligence in a machine, so machines can perform tasks that typically require human intelligence. Some programmable functions of AI systems include planning, learning, reasoning, problem solving, and decision making.

- Artificial intelligence systems are powered by algorithms, using techniques such as machine learning, deep learning and rules. Machine learning algorithms feed computer data to AI systems, using statistical techniques to enable AI systems to learn. Through machine learning, AI systems get progressively better at tasks, without having to be specifically programmed to do so.

- AI can encompass anything from Google's search algorithms, to IBM's Watson, to autonomous weapons. AI technologies have transformed the capabilities of businesses globally, enabling humans to automate previously time-consuming tasks, and gain untapped insights into their data through rapid pattern recognition. (Codebots).

What is Artificial Intelligence?

AI Defined

- AI systems are dynamic and self-improving.
- They become smarter when given more data to analyze.
- Becoming more capable and responsive as they grow.
- Effectively learning from experience.

The term AI is used a lot, and not always correctly. Many technologies exist that analyze data and produce outcomes based on that analysis, but this alone is not AI. It is not an AI system unless it can automate tasks by using cognitive abilities and reasoning. (*Asset Guardian*)

Decision-making systems, automation, and statistics are not AI.

The ability to learn and adapt is a key characteristics of an AI powered system.

What is Artificial Intelligence?

AI Systems

Types of AI systems are:
AI technologies are categorized by their capacity to mimic human characteristics, the technology they use to do this, their real-world applications, and the theory of mind

- Artificial Narrow Intelligence (ANI) – Weak form of AI – Lower than human intelligence (current AI systems)
- Artificial General Intelligence (AGI) – Strong form of AI – Human intelligence (doesn't exist, but in the works)
- Artificial Superintelligence (ASI) – Super form of AI – Surpasses human intelligence (probably the future???)

- Artificial intelligence tools are already having a significant impact on the way we conduct business worldwide, completing tasks with a speed and efficiency that wouldn't be possible for humans. However, human emotion and creativity is something incredibly special and unique, extremely difficult - if not impossible - to replicate in a machine. (Codebots).

What is Artificial Intelligence?

Subsets of AI

- Machine Learning
- Deep Learning
- Robotics
- Artificial Neural Networks
- Natural Language Processing
- Genetic Algorithms

It is not always possible to make a distinction between these categories, a single application could adopt a combination of the above categories.

What is Artificial Intelligence?

AI Theories

- One theory is based on fear of a **dystopian future**, where super intelligent killer robots take over the world, either wiping out the human race or enslaving all of humanity, as depicted in many science fiction narratives.

- The other theory predicts a more optimistic future, where humans and bots work together, humans using artificial intelligence as a tool to enhance their life experience.

What is Artificial Intelligence?

Discussion Question

Do you believe AI will take over the world and wipe out or enslave humans?

What is Artificial Intelligence?

Discussion Question

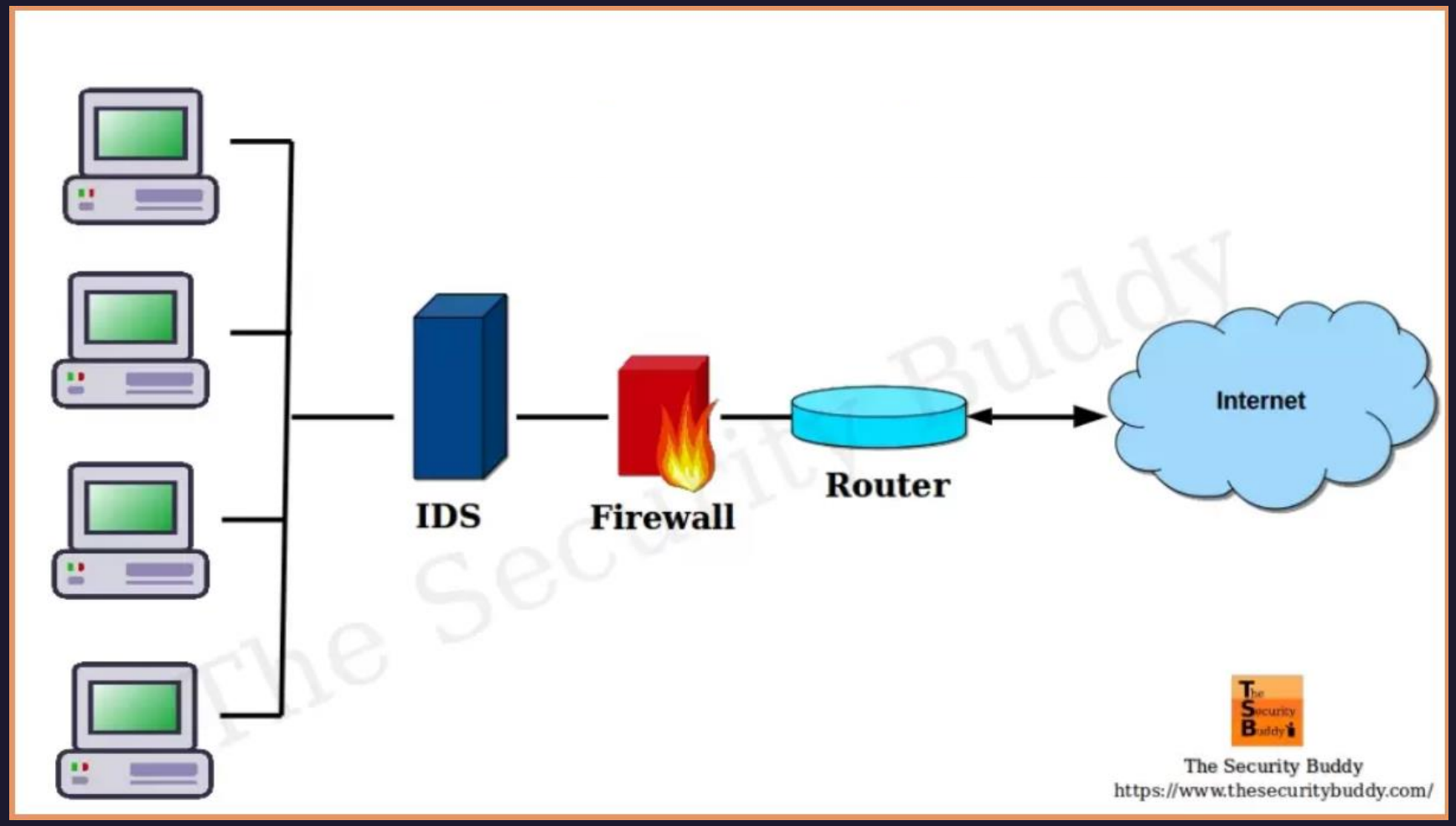
Some people believe that AI will take over many jobs in the near future. What is your take on this?

Image courtesy: ICT Mentors Solutions Limited (https://www.linkedin.com/pulse/ai-cyber-security-ict-mentors-solutions-exudf/?trk=articles_directory)

AI in Cybersecurity

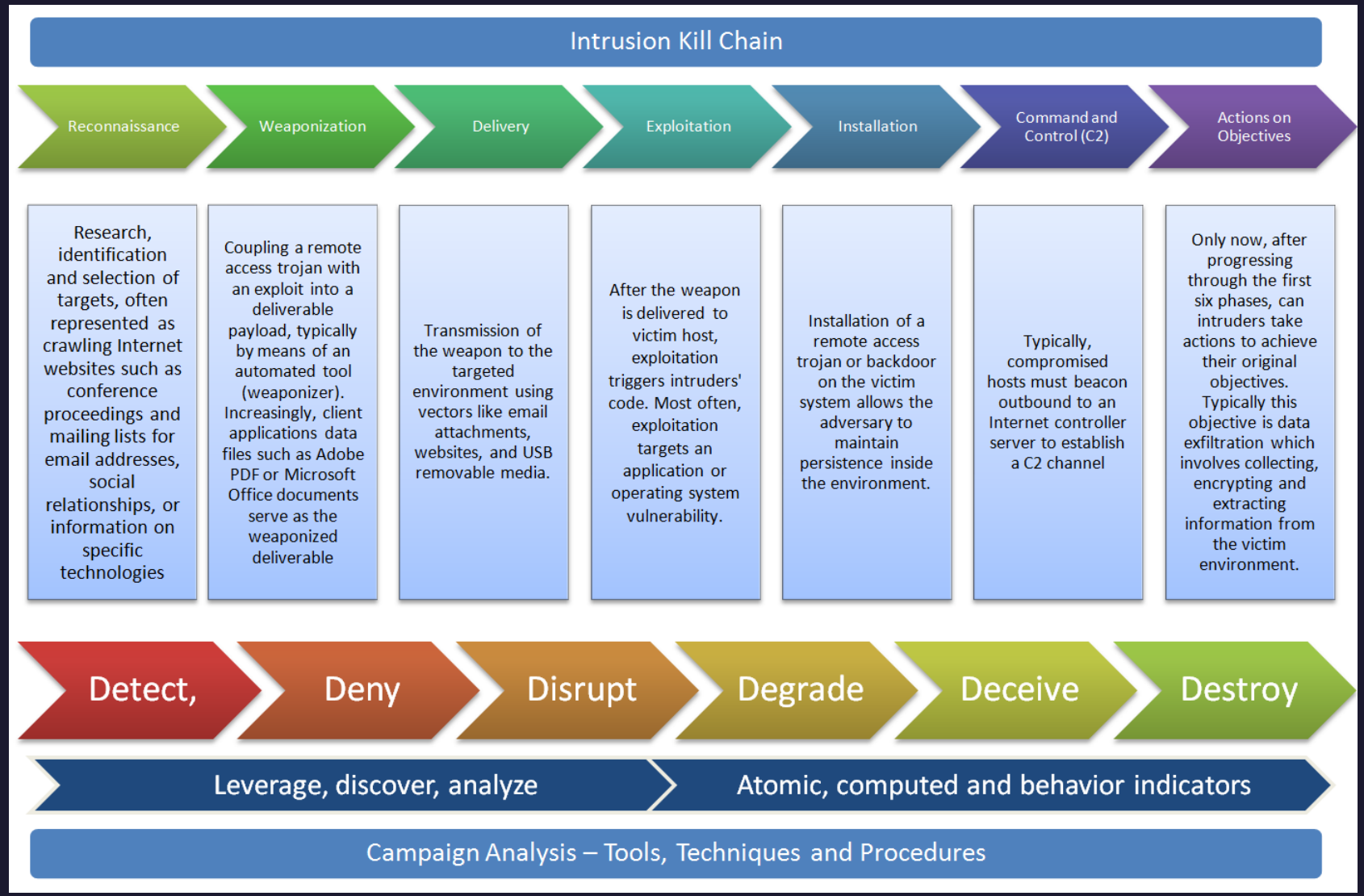
AI in Cybersecurity

A Typical Computer Network



AI in Cybersecurity

Anatomy of a Hack



AI in Cybersecurity

AAA

AAA stands for Authentication, Authorization, and Accounting (sometimes called Auditing). AAA are elements of Access Control which is also a goal of security.

- **Authentication**: The process of verifying a user, device, or process before allowing access to a system or resources
- **Authorization**: This is the set of permissions given to a user, device, or process to access resources after a successful authentication.
- **Accounting**: This is the process of monitoring user activities while they are logged in to a system. This includes recording user activities in logs (logging) and reviewing the logs after the fact (auditing). Accounting is dependent on successful authentication.

The entire process of system use generates and stores lots of data. In an enterprise environment, this huge data needs to be reviewed continuously to ensure anomalous activities are not missed.

AI in Cybersecurity

Data! Data!! Data!!! Everywhere

- Approximately **328.77 million terabytes of data are created each day**
- Around **120 zettabytes of data will be generated this year**
- **181 zettabytes of data will be generated in 2025**
  (Exploding Topics)

- Connected devices are projected to generate a staggering **79 zettabytes of data by 2025**, manual analysis by humans becomes impractical, making AI an indispensable tool in the fight against cybercrime (TechMagic).

AI in Cybersecurity

Data! Data!! Data!!! Everywhere

Abbreviation	Unit	Value	Size (in bytes)
b	bit	0 or 1	1/8 of a byte
B	bytes	8 bits	1 byte
KB	kilobytes	1,000 bytes	1,000 bytes
MB	megabyte	$1,000^2$ bytes	1,000,000 bytes
GB	gigabyte	$1,000^3$ bytes	1,000,000,000 bytes
TB	terabyte	$1,000^4$ bytes	1,000,000,000,000 bytes
PB	petabyte	$1,000^5$ bytes	1,000,000,000,000,000 bytes
EB	exabyte	$1,000^6$ bytes	1,000,000,000,000,000,000 bytes
ZB	zettabyte	$1,000^7$ bytes	1,000,000,000,000,000,000,000 bytes
YB	yottabyte	$1,000^8$ bytes	1,000,000,000,000,000,000,000,000 bytes

World Economic Forum

AI in Cybersecurity

Data! Data!! Data!!! Everywhere

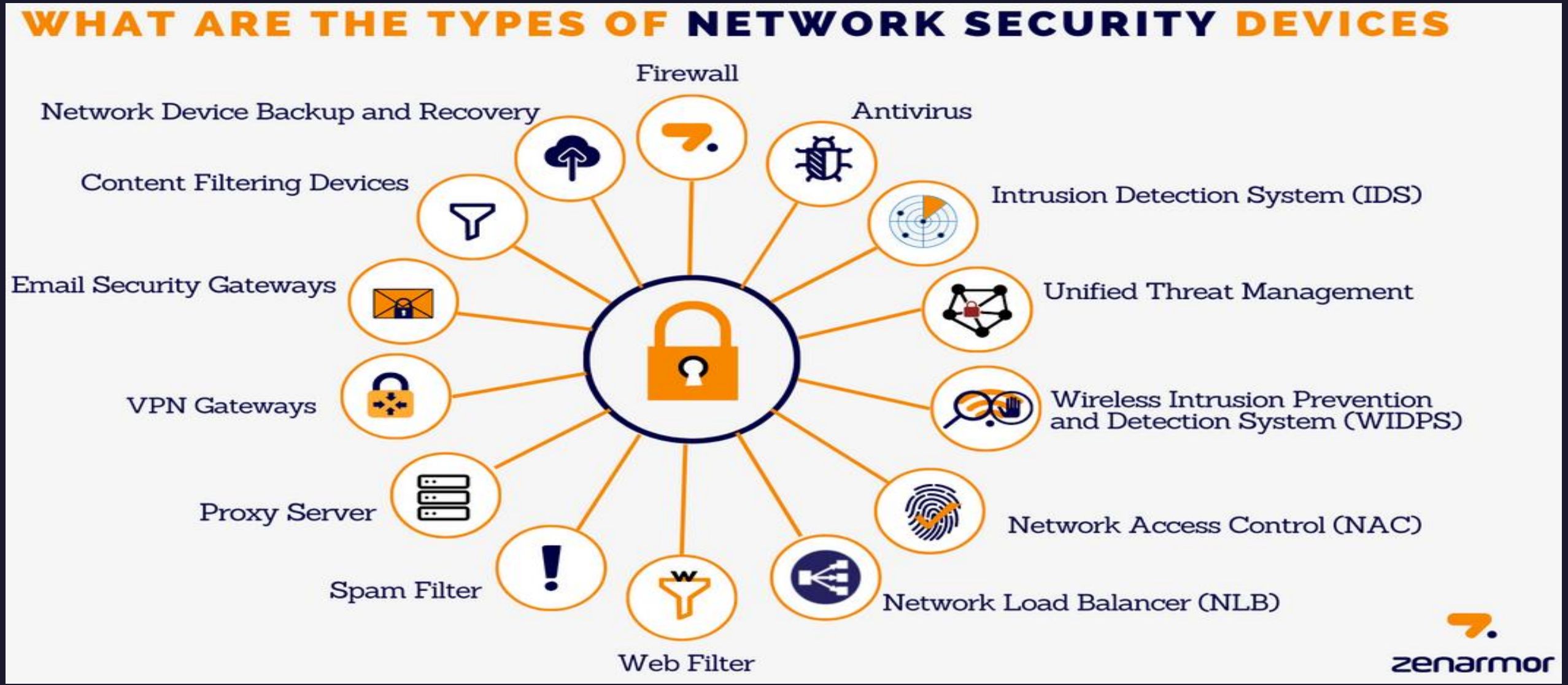
One Gigabyte Of Data

- **Text files:** Nearly 678,000 pages per gigabyte.
- **Emails:** More than 100,000 pages.
- **Microsoft Word files:** Almost 65,000 pages.
- **PowerPoint Slide Decks:** Roughly 17,500 pages.
- **Images:** Close to 15,500 pages.

- One zettabyte is approximately equal to *1,000 exabytes or 1 billion terabytes or 1 trillion gigabytes*

AI in Cybersecurity

Network Security Devices



AI in Cybersecurity

Network Security Devices

AI is used in cybersecurity to create security applications and devices like the following:

- Antivirus Software
- Intrusion Detection Systems (IDS)
- Intrusion Prevention Systems (IPS)
- Security Information and Event Management
- Security Orchestration Automation and Response (SOAR)
- Firewalls (Network/Host)
- Next-Generation Firewalls (NGFW)
- Unified Threat Management (UTM)

AI in Cybersecurity

Threat Landscape

- A broad attack surface
- Hundreds of devices to protect in each organization
- Hundreds of attack vectors that cybercriminals can exploit
- A significant shortage of skilled security professionals to handle the growing demands
- Massive amounts of data that have surpassed human-scale processing capacity, making it a daunting task to analyze and make sense of.

According to TechRepublic, mid-sized companies receive over **200,000** alerts for cyber events each day, and a team of human experts cannot possibly address all of them. Consequently, certain threats are likely to go unnoticed, leading to significant network damage. To overcome these challenges, businesses seeking to succeed in the digital world must rely on AI and other advanced technologies to bolster their cybersecurity defenses. (Engati)

AI in Cybersecurity

Use Cases of AI in Cyber Security

- **Threat Detection and Prevention**
  - Malware and Phishing Detection
  - Security Log Analysis
  - Endpoint Security

- **User Behavior Analytics**
- **Advanced Threat Response and Mitigation**
- **Vulnerability Assessment and Management**
- **Security Operations and Automation**
- **Threat Intelligence and Predictive Analytics**

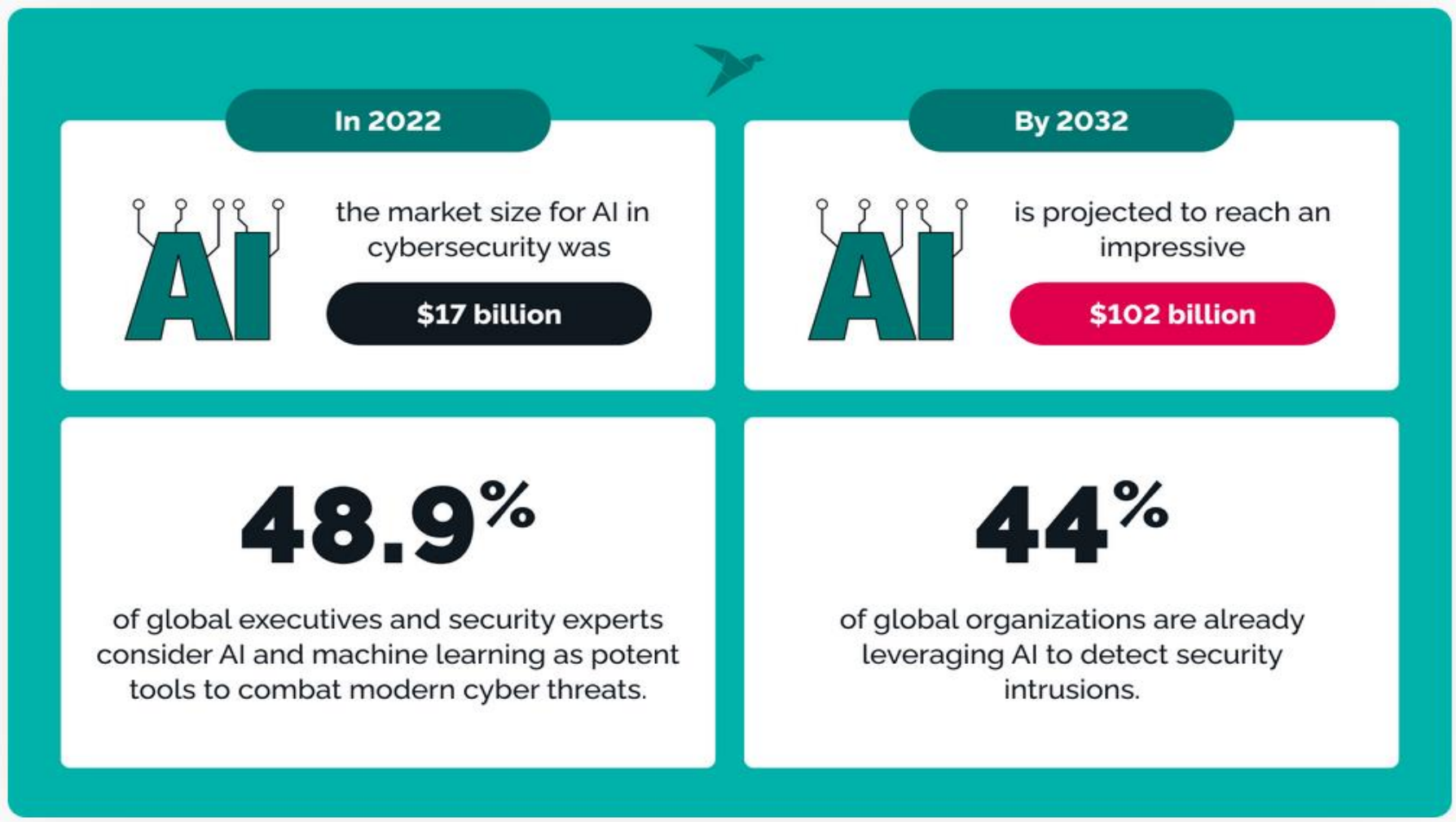
AI in Cybersecurity

Current Applications

- Breach risk prediction
- Phishing detection
- Malware detection & prevention
- User authentication
- Spam filtering
- Password protection
- Bot identification
- Behavioral analysis
- Network segmentation & security
- Fraud detection
- Threat intelligence
- Incident response
- Vulnerability management
- Identity & access management

AI in Cybersecurity

The Numbers



(TechMagic)

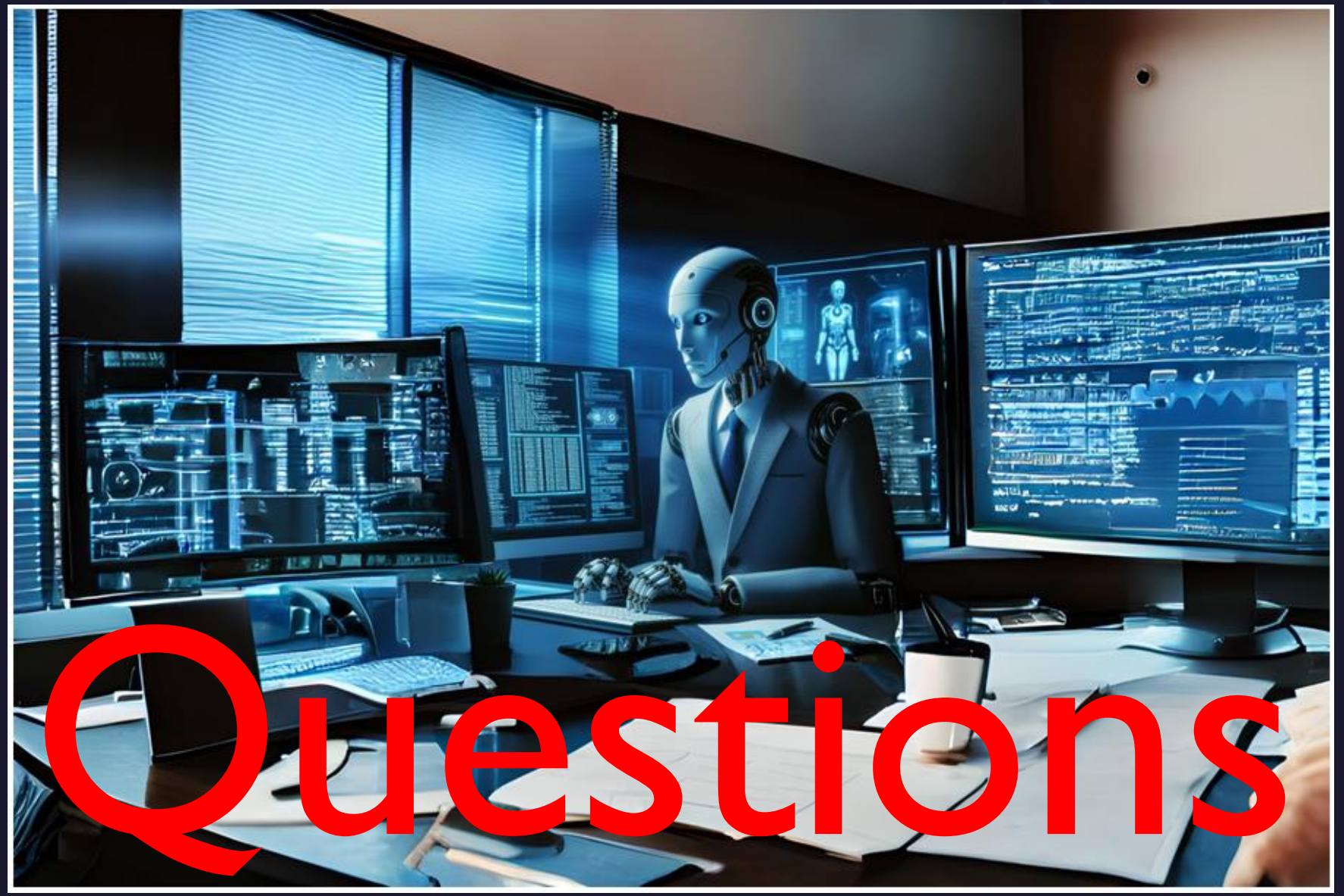


Image courtesy: ICT Mentors Solutions Limited (https://www.linkedin.com/pulse/ai-cyber-security-ict-mentors-solutions-exudf/?trk=articles_directory)

AI in Cybersecurity

The Bad

- Artificial Intelligence (AI) can be considered a double agent in terms of its role in Cyber Security. While it can be used to protect networks and systems it can also be used to attack them.

- Malicious AI can be used to identify patterns and weaknesses in Cyber Security systems and then exploit them. Companies will have to evolve and utilize the same AI techniques to counter these attacks.

- Phishing emails that have been created by AI help attackers target and convince victims that the communication is genuine. This particularly helps attackers who want to send emails to people in other countries where they do not speak the local language

- Malware that has an AI component may change and adapt to avoid detection. It may then lurk inside a system, gathering data and observing users, and either stealthily sending the data back to the attackers or waiting until the most opportune time to launch another form of attack.

AI in Cybersecurity

The Bad

- AI systems such as ChatGPT and Google Bard use deep learning to generate realistic text from large amounts of data:

- They can create
  - Chatbots
  - Stories
  - Code

- Misused to create
  - Fake News
  - False Information
  - False Identity
  - Scam Emails

AI in Cybersecurity

The Bad

- Large Language Models (LLMs) such as ChatGPT and Google Bard (Gemini) now have evil cousins that are dedicated to exploiting computer systems and networks or spreading fake information.
  - PoisonGPT
- There's good and bad information online, AI systems trained on bad information will spread such information.
- Also, AI systems designed to exploit systems will do just that.
- AI systems are not perfect !!!
- Trust but verify.

AI in Cybersecurity

The Bad

Malicious AI systems discovered recently:

- Evil-GPT
- WormGPT
- FraudGPT
- XXXGPT
- Wolf GPT

(Cyber Security News)

AI in Cybersecurity

The Bad

AI Powered Malware

AI in Cybersecurity

The Bad

https://youtu.be/VzPEx0mlePQ

AI Assisted Hacking

AI in Cybersecurity

Deep Fake

The Bad

AI in Cybersecurity

The Bad

Voice Cloning Phone Scam

https://youtu.be/3g0UC3jpvzU

AI in Cybersecurity

The Bad

The making of a Voice Cloning Phone Scam

AI in Cybersecurity

The Bad

AI in Elections

AI in Cybersecurity

The Bad

AI in Elections

AI in Cybersecurity

The Good

Increasingly companies cannot rely on humans alone to deal with the amount of data generated and the complexity of their Cyber Security needs. Artificial Intelligence (AI) is now used by many businesses to manage infrastructure, interpret data, and automatically counter cyberattacks. (Asset Guardian)

- Voluminous and repetitive tasks
- AI learns more over time
- Detecting and managing known threats
- Identifies unknown threats
- Processing large quantities of data
- Handling vulnerability and real-time security events
- Quicker response time compared to humans
- Phishing
- Better Overall Security

- Securing Authentication
- Eliminates time-consuming tasks
- Immune to fatigue
- Email filters and warnings,
- Malware identification and threat detection.
- Handling bots
- Reduces Duplication of processes
- Improved Accuracy
- AI is not yet able to break encryption

AI in Cybersecurity

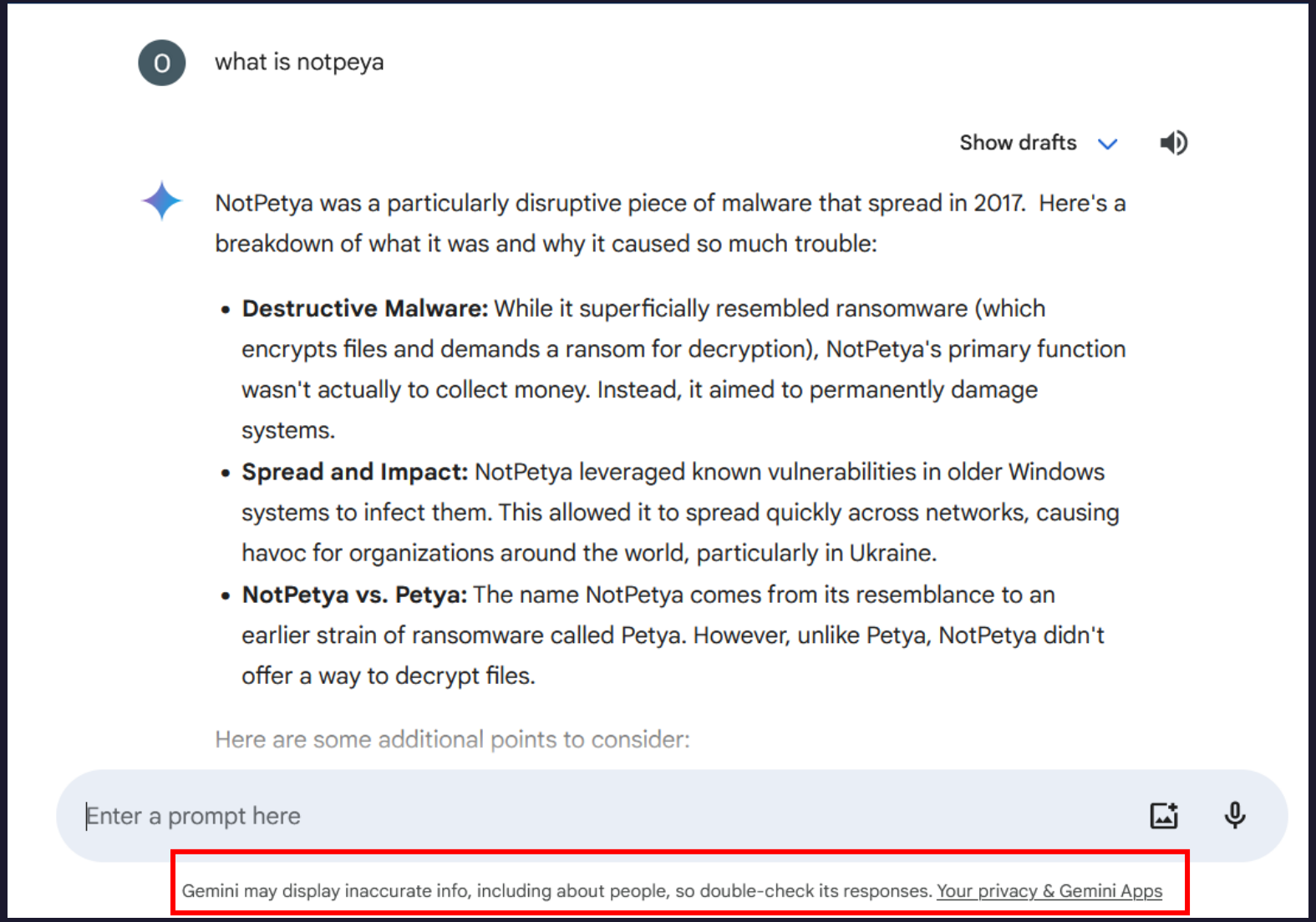
The Good

- **Investigate**: Investigate incidents, suspicious files, APT campaigns, etc.
- **Identify**: Specific artifacts, elements, Groups, nation state actors.
- **Report**: Report past attacks and future threats.
- **Research**: Conduct deeper research about relevant security subjects.

AI in Cybersecurity

The Good

AI for cybersecurity research (Gemini)



AI in Cybersecurity

The Good

https://youtu.be/mpw-olvjB70

AI for Physical Security

AI in Cybersecurity

The Good

In IBM's 2022 data security report, it was reported that it took an average of 277 days – **roughly 9 months** – for businesses to identify and report a data breach. Stolen or compromised credentials were the most common cause of a data breach in 2022, and these types of attacks took around 327 days to identify. (CYFORSECURE Cyber Security)

According to IBM, AI can slash the time taken to detect and respond to cyber threats by a staggering 14 weeks (TechMagic)

AI in Cybersecurity

The In-Between

The risks of AI are not limited to its application as a form of attack, there are some risks associated with using it in any form.

- AI may not work well for smaller companies that have smaller datasets, and there will always be concerns around privacy and data retention laws.
- There must be a balance between keeping the data anonymous but at the same time remaining usable.
- Bias in AI security systems
- Overreliance on AI systems may affect normal security practices.
- Misinterpretation (AI hallucinations)
- There is a shortage of Cyber Security AI/Machine Learning professionals, with the demand increasing as systems become more complex.

AI in Cybersecurity

The In-Between

- It can be very expensive to implement and may not be a viable solution for smaller businesses
- AI systems are not infallible and may be tricked into incorrect behavior where more rigid systems would not be.
- As we think of AI in Cyber Security, we also need to think of Cyber Security for AI.
- AI systems can be as vulnerable to attack as any other system and AI is only as clever as the data that is fed into it.
- By manipulating the data, attackers may be able to trick the AI into behaving against its intended design, giving false positives or bypassing security.
- The protection of AI from attack is still a new concept, but policies and standards are being developed by organizations like the Brookings Institution and the ETSI Industry Specification Group on Securing Artificial Intelligence.

AI in Cybersecurity

The In-Between

AI does a lot of good, but there are some inherent dangers with AI as well.

- Should we be concerned? Maybe! While a chatbot will not tell you "How to build a bomb" there are ways to bypass its protections, the data set for chatbots still remains the entire internet.
- By breaking up the creation of malicious software into smaller more innocent parts, chatbots will produce the various components.
- ChatGPT for instance is designed to be extremely cautious about creating or revealing anything that is straightforwardly malicious or unethical. However, it is not as cunning as a human trying to social engineer it into doing so. For instance, by treating a situation as hypothetical or fictional the AI will reveal or create more than if it had been a straight question
- As with all new technologies the regulators are racing to catch up. An unusual aspect of regulating AI is the "black box" nature of the code.
- While software companies can be secretive about how their systems work, they do at least know how it works.
- With AI systems though, how the code works is a mystery even to those that developed the AI, which poses an interesting problem for those that wish to impose policies and regulations on it.
- How much of this is fearmongering though, remains to be seen (Asset Guardian).

AI in Cybersecurity

The Future

- Next step for AI is Artificial General Intelligence (AGI), the type of AI that can understand and learn as well as any human.
- In the nearer future, AI will take over more tasks from humans, according to Gary Kasparov, "Jobs don't disappear, they evolve."
- Eventually, there will be more laws, to regulate AI use, just like all aspects of technology and life.
- There will be more chatbots such as ChatGPT and Gemini, they will do good, they may also do some harm, but they are here to stay – they are the future.
- Chatbots and AI in general will be used to create cyberattacks, AI will also be used to combat cyber attacks as well.
- Artificial Intelligence already plays a significant part in Cyber Security, and on its current path will take over more tasks and decision-making from humans. It will be a long time before it is smart enough to do everything unattended, but with each new technological breakthrough, we come closer to that possibility (Asset Guardian).

AI in Cybersecurity

Discussion Question

Is AI good or bad for cybersecurity?

Summary

AI is here to stay, we will all have to learn to live with it as much as we can. AI in cybersecurity is here to stay as well, there is no going back.

Thank You

O. Cheche Agada (Ph.D.)

oagada@gmu.edu

References

- Asset Guardian (https://www.assetguardian.com/how-ai-is-changing-the-cyber-security-landscape/)
- Codebots (https://codebots.com/artificial-intelligence/the-3-types-of-ai-is-the-third-even-possible)
- The Conversation (https://theconversation.com/not-everything-we-call-ai-is-actually-artificial-intelligence-heres-what-you-need-to-know-196732)
- Akkio (https://www.akkio.com/post/the-five-main-subsets-of-ai-machine-learning-nlp-and-more)
- Count Upon Security (https://countuponsecurity.com/2014/08/29/intelligence-driven-incident-response/)
- Engati (https://www.engati.com/blog/ai-for-cybersecurity)
- TechMagic (https://www.techmagic.co/blog/ai-in-cybersecurity/)

CYFORSECURE CYBER SECURITY (https://cyforsecure.co.uk/how-long-does-it-take-to-detect-a-cyber-attack/)

References

- Exploding Topics (https://explodingtopics.com/blog/data-generated-per-day)
- World Economic Forum (https://www.weforum.org/agenda/2019/04/how-much-data-is-generated-each-day-cf4bddf29f/)
- Cyber Security News (https://cybersecuritynews.com/hackers-released-evil-gpt/)