AMITT

ADVERSARIAL Misinformation Influences tactics and technique

Prepare, Spread, Detect, Mitigate, technique.

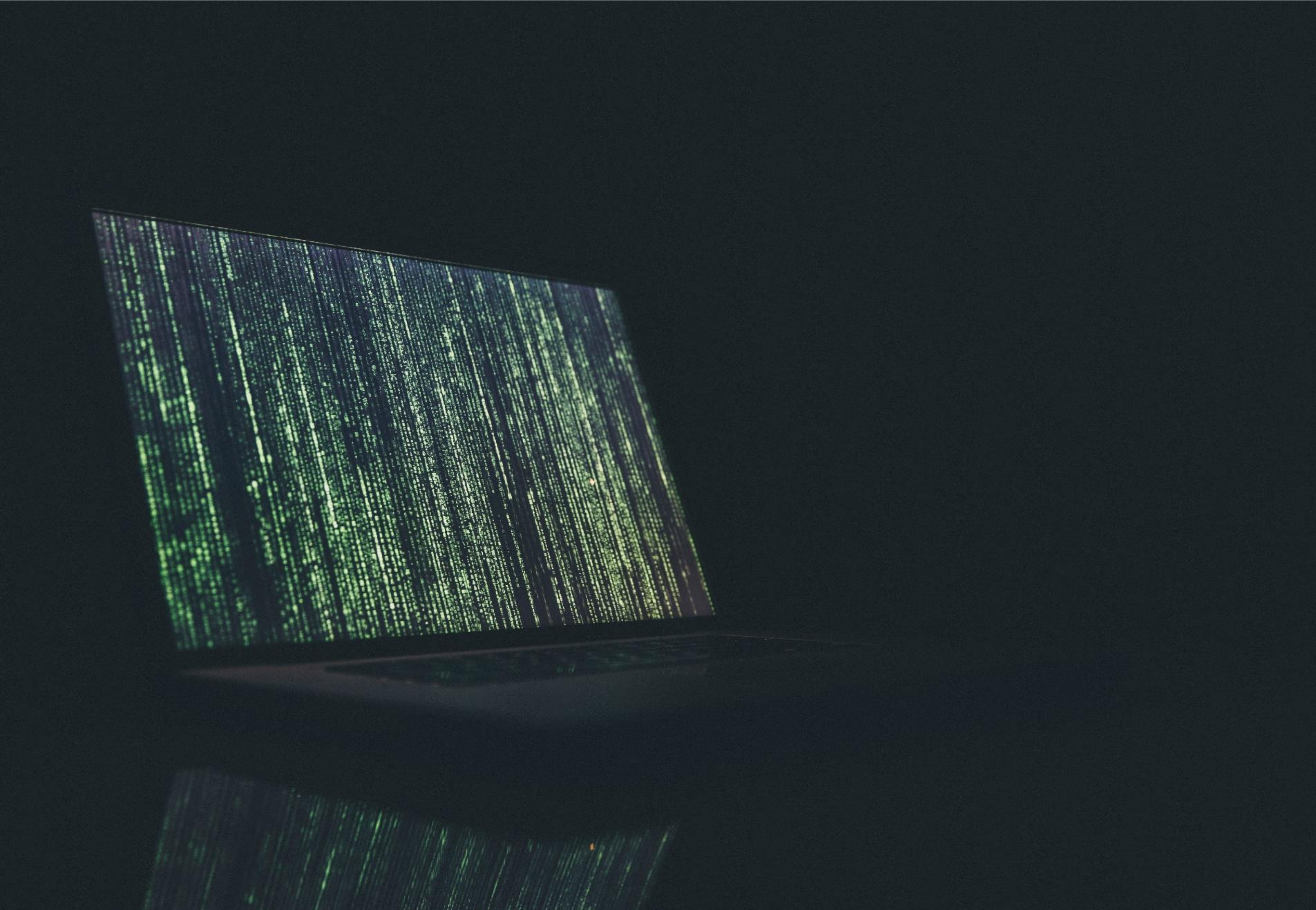


Table of Contents

Introduction

AMITT, which stands for Adversarial Misinformation Influences Tactics and Techniques, is a set of standards and an open-source knowledge base of both red-team and blue-team disinformation tactics and techniques. AMITT has a purpose that people respond to disinformation incidents, plan defenses and countermoves, and transfer information security principles to the disinformation sphere. it provides a common taxonomy of cognitive security offense and defense, a framework for rapidly sharing threat intelligence, and a conceptual tool for strengthening disinformation defenses through red teaming, risk analysis, replays, and simulations.

With its comprehensive approach, AMITT serves as a valuable tool in the field of information integrity. It empowers stakeholders to better identify, analyze, and respond to the challenges posed by disinformation, ultimately promoting a more informed and resilient society.

Digital Disinformation

Disinformation is digital harm. A group of people Intentionally spread false, misleading, or misattribute information to achieve their goals by changing human beliefs or emotions in many people.

Effects of Disinformation Include:

- **Physical harm:** Damage to infrastructure, Hardware, Body injury
- **Psychological harm**: Anxiety, Depression due to cyber bullying, cyber stalking, etc.
- **Economic harm:** Financial loss e.g., from data breach cybercrime etc.
- **Reputational harm:** organizations' loss of consumers, individuals' disruption of personal life, country s damage trade negotiations
- **Cultural harm**: Increase in social disruption creating real-world violence.
- **Political harm:** Disturbance in the political process, lost government services from Internet shutdown, botnets influence votes.

Motivation behind

- **Harm the reputation of an organization, people, or country**: Attackers spread false or misleading information about the target to damage their reputation in the public eye.
- **Money:** Attackers make money by No of view the website, clicking on something, revenue generating, selling disinformation, selling merchandise, and selling accounts(botnet).
- **Geopolitics**: Countries use disinformation to change external opinion about them.
- **Politics and power:** Disinformation can be used to gain or maintain political power and also create, subvert, or hijack political and activist groups.
- **Business:** Disinformation can be used to harm the business deal and partnership, PR and advertisement, R&D and capital investment.
- **Attention and fun: P**eople spread disinformation for fun, popularity, and make money.

Influence Techniques

Influence techniques are strategies used to manipulate or shape the opinions, beliefs, and behaviors of individuals or populations. Some common influence techniques include:

- The four Ds: Distort, Distract, Dismay, Dismiss
  - Distort: Manipulating information or facts to mislead or misrepresent the truth.
  - Distract: Introducing irrelevant or sensationalized topics to divert attention from important issues.
  - Dismay: Evoking negative emotions such as fear, anger, or sadness to manipulate public opinion.
  - Dismiss: Discrediting valid information or opposing viewpoints to undermine their credibility.

  These four Ds represent common tactics used in influencing public opinion. Understanding and recognizing these techniques can help individuals critically analyze information and make informed decisions.

- Hack and leak are a strategy where unauthorized access to sensitive information is obtained through hacking, and then that information is publicly released or leaked. It aims to expose hidden data, manipulate public opinion, or raise awareness. It is illegal and unethical.
- Unsurmountable proof refers to the use of compelling and indisputable evidence to influence others. It involves presenting irrefutable facts or data that make it extremely difficult for others to challenge or dispute, aiming to persuade by appealing to logic and credibility.

Fake (and real) content

- **Misinformation**: Misinformation is false or misleading information that is spread unintentionally. It can be spread through a variety of channels, such as social media, news websites, and word-of-mouth. Misinformation can be harmful because it can lead people to make decisions that are not in their best interests.
- **Fake but credible research:** Fake but credible research is research that is published in a legitimate academic journal or other publication, but which is false or misleading. This type of research can be harmful because it can give people the impression that a particular claim is true when it is not.
- **Grain of truth:** (Thumb rule 90% true information, 10% misinformation) Grain of truth is information that is partially true, but which is presented misleadingly. For example, a grain of truth story might say that "vaccines cause autism" when in fact there is no scientific evidence to support this claim.
- **Narratives:**(Storytelling “in-group” and “out-group”) Narratives are stories that are told to explain or justify a particular event or situation. Narratives can be used to spread misinformation by presenting false or misleading information in a way that is emotionally appealing or persuasive. For example, a narrative might say that "the government is trying to control us" when in fact there is no evidence to support this claim.

Fake (and real) Accounts

  - Bots: The bulk of malicious accounts are handled by one attacker, Automated to spread Disinformation.
  - Parody: Clearly counterfeit account used to satirize or diminish image.
  - Spoof: Counterfeit account which closely copies a
  - real account.
  - Deep Cover: False account accepted as a real account for a long period.
  - Take Cover: Real account controlled by someone who isn't its owner.
  - Camouflage: False account which is a mimic of the community of real accounts.
- **Ignorant agent (independent agent):** These are people who spread disinformation because they believe it to be true. They may not be aware that the information they are sharing is false, or they may be unable to verify its accuracy.
- **Fake expert (story planted)**: These are people who are paid to spread disinformation. They may be experts in a particular field, or they may simply be good at creating convincing stories. Their goal is to plant false information in the media or online so that it will be seen by a large number of people.
- **Fake groups (computational propaganda):** These are groups of people who are organized to spread disinformation online. They may use social media, forums, or other online platforms to spread their messages. Their goal is to sow discord and division among the public.
- **Disinformation websites (advertisement money):** These are websites that are created with the sole purpose of spreading disinformation. They may make money through advertising, or they may be funded by organizations that have a vested interest in spreading false information.
- **Pink slime network (Misinformation spread through the local newspaper):** This is a type of disinformation that is spread through local newspapers. The goal is to create the illusion of widespread support for a particular viewpoint or damage the reputation of a particular person or organization.

Fake (and real) sharing

- **Amplification (Through groups):** This is the act of sharing information or content within a group of people, such as a social media group or a chat group. This can be done to spread information quickly and widely or to create the illusion of widespread support for a particular viewpoint.
- **Hashtag jacking:** This is the act of using a popular hashtag to spread disinformation. This can be done by creating posts or tweets that use the hashtag, or by commenting on posts that use the hashtag. The goal of hashtag jacking is to get the disinformation seen by as many people as possible.
- **Microtargeting:** This is the act of targeting specific individuals or groups with disinformation. This can be done by using data about people's interests, demographics, or online behavior. The goal of microtargeting is to make disinformation more believable and persuasive to the people who see it.
- **Search Engine Optimization (Black-hat SEO):** This is the act of manipulating search engine results to make disinformation more visible. This can be done by using keywords and phrases that are associated with disinformation, or by creating fake websites that rank high in search results. The goal of black-hat SEO is to make it easier for people to find disinformation.

- **Manipulate online polls:** This is the act of voting in online polls in a way that is designed to skew the results. This can be done by using bots or by manually voting multiple times. The goal of manipulating online polls is to make it appear that there is more support for a particular viewpoint than there is.
- **Organizing real-world events:** This is the act of organizing real-world events that are designed to spread disinformation. This can be done by holding rallies, protests, or other events. The goal of organizing real-world events is to give disinformation a more tangible presence and to make it more persuasive.
- **Astroturfing and information pollution (firehouse of misinformation):** This is the act of creating and spreading large amounts of disinformation in a way that is designed to overwhelm and confuse people. This can be done by using social media, websites, or other online platforms. The goal of astroturfing and information pollution is to make it difficult for people to know what is true and what is false.

Cognitive security

Cognitive security systems are designed to understand how humans think and behave and to use this understanding to identify and defend against cyberattacks. They can do this by analyzing large amounts of data, such as social media posts, email messages, and website traffic, to identify patterns that may indicate a cyberattack. They can also use AI to generate realistic-looking phishing emails and other social engineering attacks, to train users to spot them.

Here are some of the benefits of cognitive security:

- It can help to identify and defend against new and emerging threats. Cognitive security systems are constantly learning and evolving, so they can keep up with the lates**t cyberattacks.**
- **It can be more effective than traditional security methods. Cognitive security systems can use AI to analyze large amounts of data and identify patterns that may indicate a cyberattack. This can help to prevent attacks before they happen.**
- **It can be more efficient than traditional security methods. Cognitive security systems can automate many of the tasks that are currently done by human security analysts. This can free up security analysts to focus on more complex tasks.**
- **It can be more cost-effective than traditional security methods. Cognitive security systems can be more cost-effective than tra**ditional security methods, such as firewalls and intrusion detection systems. This is because they can be used to protect a wider range of threats and can be more effective at preventing attacks.

Cognitive Security Threat Response

- **Identifying and prioritizing threats:** Cognitive security systems can be used to identify and prioritize threats based on a variety of factors, such as the likelihood of the threat being successful, the potential damage that the threat could cause, and the resources that are available to respond to the threat.
- **Mitigating threats:** Cognitive security systems can be used to mitigate threats by taking a variety of steps, such as blocking malicious traffic, removing malware, and educating users about security risks.

- **Responding to incidents:** Cognitive security systems can be used to respond to incidents by taking a variety of steps, such as tracing the source of the attack, notifying affected users, and restoring systems to a previous state.
- **Recovering from incidents:** Cognitive security systems can be used to recover from incidents by taking a variety of steps, such as restoring systems to a previous state, implementing new security controls, and training users about security risks.

Cognitive Security Services

- **Social media monitoring**: This service uses AI to monitor social media for signs of disinformation. It can identify and flag posts that are likely to be false or misleading, and it can also track the spread of disinformation across social media platforms.
- **Fact-checking:** This service uses AI to verify the accuracy of claims made online. It can cross-reference claims with other sources of information, and it can also use natural language processing to identify patterns that may indicate false or misleading information.
- **Digital forensics:** This service uses AI to investigate cyberattacks. It can analyze data from compromised systems to identify the source of the attack, and it can also track the attacker's movements.
- **Threat intelligence:** This service uses AI to collect and analyze threat data. It can identify emerging threats, and it can also track the activities of known threat actors.
- **Training:** This service uses AI to train users to spot disinformation. It can generate realistic-looking phishing emails and other social engineering attacks, to train users to identify them.

Tools for Detection

- Google's Fact Check Explorer: Fact Check Tools Recents (google.com)
- Snopes: http://www.snopes.com/
- NewGuard: http://www.newsguardtech.com/
- Tenderizer: Trendolizer™ - Trendolizer

Practical

Attack Technique 01: Disinformation for Attention and Fun

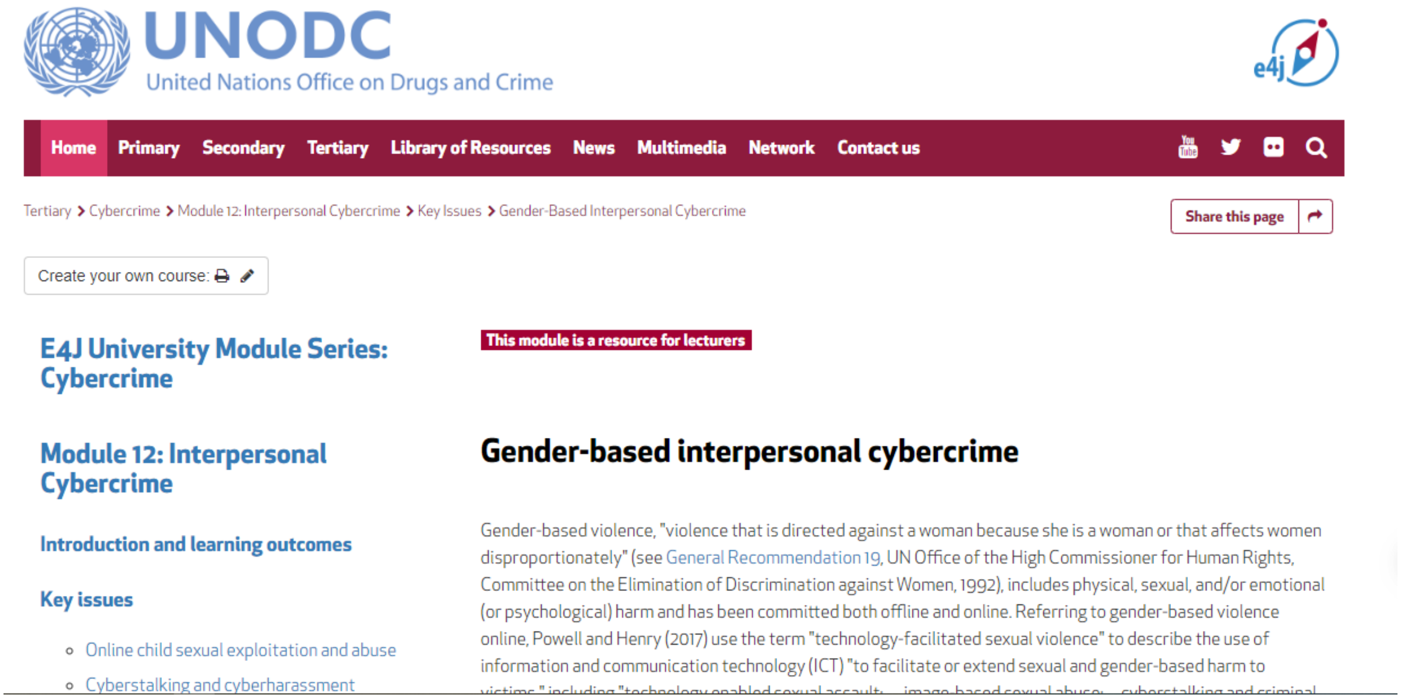
Disinformation for attention and fun refers to the deliberate creation and dissemination of false or misleading information with the primary goal of gaining attention, amusement, or causing disruption. It typically involves spreading rumors, hoaxes, or fabricated stories through various mediums, such as social media platforms, websites, or online forums.

The intention behind disinformation for attention and fun is to deceive and manipulate others, often for personal satisfaction or to provoke a reaction. However, it's important to note that engaging in such activities can have serious consequences. Disinformation can lead to confusion, harm innocent individuals, damage reputations, erode trust in institutions, and even incite social or political unrest.

Disinformation campaigns can range from spreading false rumors about individuals or organizations to creating fabricated news stories, manipulating images or videos, or even launching coordinated disinformation campaigns to influence public opinion. These campaigns often exploit people's emotions, biases, or vulnerabilities to achieve their objectives.

It's crucial to prioritize ethical behavior online and promote responsible use of the internet. Instead of participating in disinformation campaigns, it's advisable to focus on positive and constructive contributions to society and engage in activities that foster genuine dialogue, knowledge sharing, and critical thinking.

Let's say an individual wants to gain attention and create chaos on social media. They create a fake news story claiming that a popular celebrity has been involved in a scandalous incident. They craft a convincing narrative, complete with fabricated evidence, doctored images, and fake witness accounts. They then release the story on a website they have created and share it on various social media platforms using multiple fake accounts.


UNODC
United Nations Office on Drugs and Crime
e4j
Home
Primary
Secondary
Tertiary
Library of Resources
News
Multimedia
Network
Contact us
Tertiary > Cybercrime > Module 12: Interpersonal Cybercrime > Key Issues > Gender-Based Interpersonal Cybercrime
Share this page
Create your own course:
E4J University Module Series: Cybercrime
Module 12: Interpersonal Cybercrime
Introduction and learning outcomes
Key issues
Online child sexual exploitation and abuse
Cyberstalking and cyberharassment
This module is a resource for lecturers

Gender-based interpersonal cybercrime

Gender-based violence, "violence that is directed against a woman because she is a woman or that affects women disproportionately" (see General Recommendation 19, UN Office of the High Commissioner for Human Rights, Committee on the Elimination of Discrimination against Women, 1992), includes physical, sexual, and/or emotional (or psychological) harm and has been committed both offline and online. Referring to gender-based violence online, Powell and Henry (2017) use the term "technology-facilitated sexual violence" to describe the use of information and communication technology (ICT) "to facilitate or extend sexual and gender-based harm to victims," including "technology-enabled sexual assault; image-based sexual abuse; cyberstalking and criminal

This figure shows a website that is purely operated by UNODC for reporting such cyber humiliations.

Case Study:

During the 2016 U.S. presidential election, there was a cybercrime incident where Russian operatives spread false information on social media to influence voters. They created fake accounts and groups to share misleading content that deepened political divisions and caused confusion. Their goal was to sway public opinion and disrupt the election process. This incident led to investigations and concerns about the manipulation of information online, leading to calls for better regulation and media literacy to combat disinformation.

Russia 'meddled in all big social media' around US election

17 December 2018

GETTY IMAGES

The report suggests YouTube, Tumblr, Pinterest, Instagram and Google+ were affected by Russian interference, as well as Facebook and Twitter

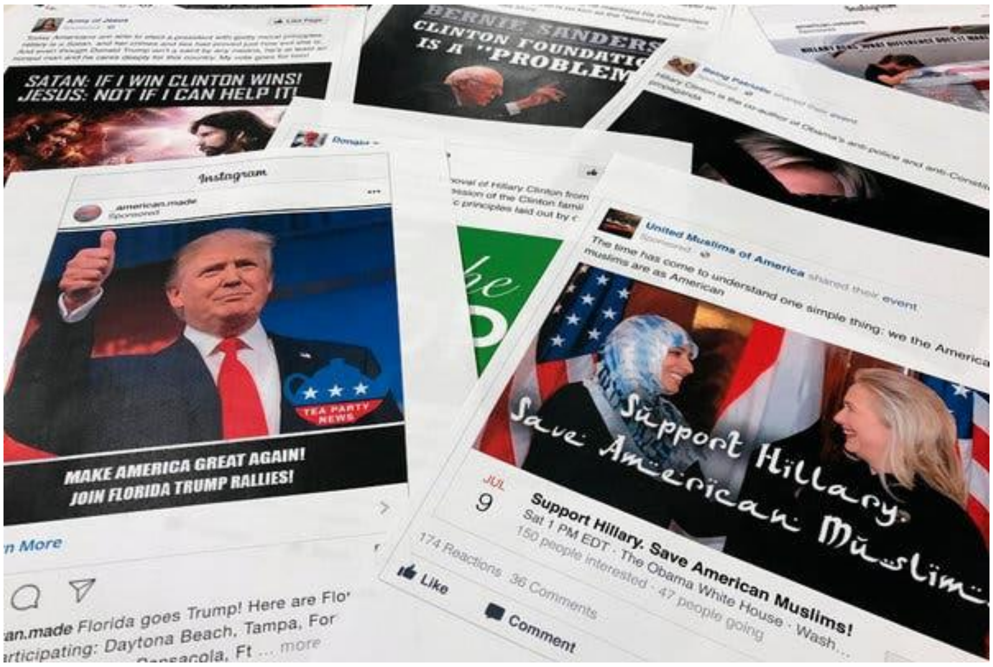
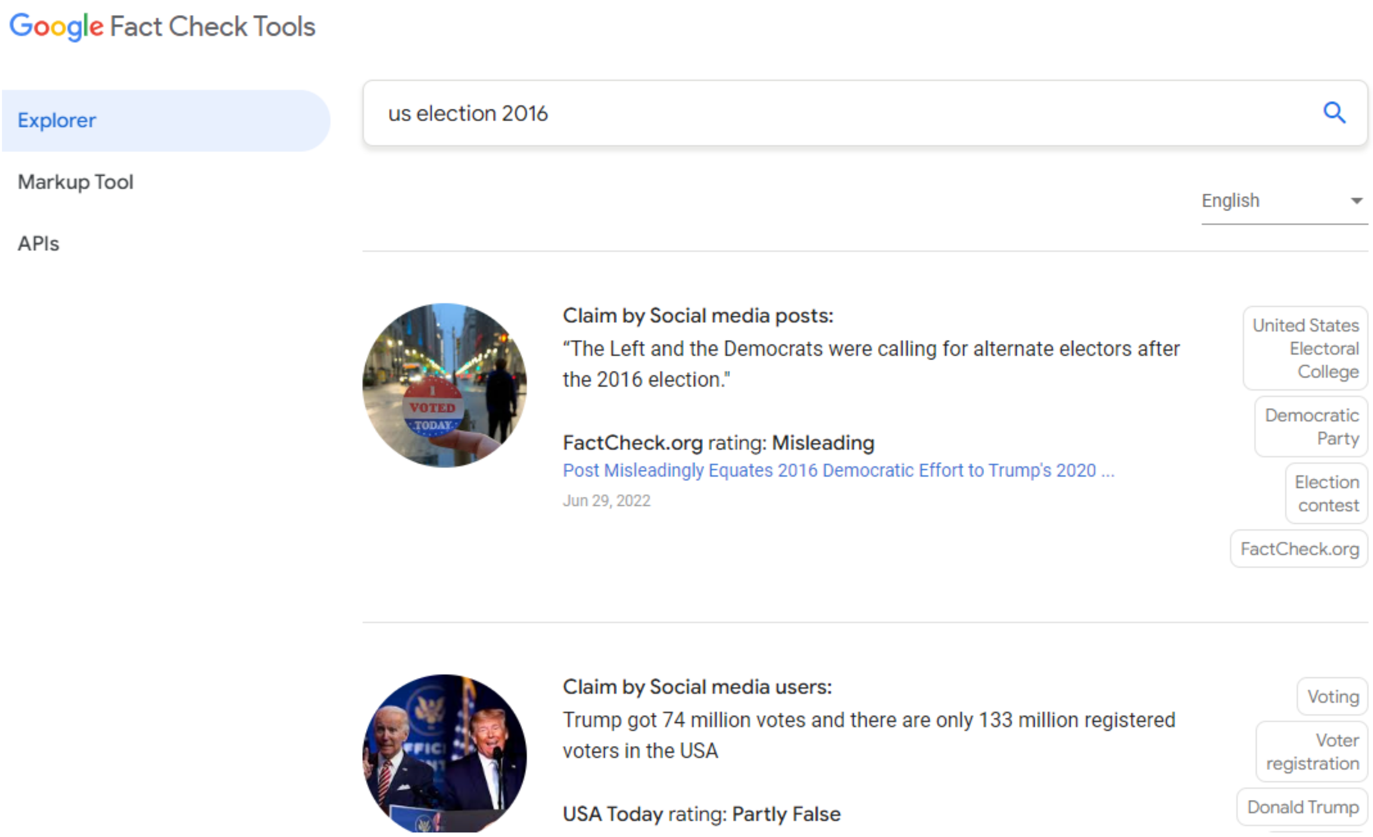

SATAN: IF I WIN CLINTON WINS!
JESUS: NOT IF I CAN HELP IT!
BERNIE SANDERS:
CLINTON FOUNDATION IS A "PROBLEM"
Instagram
americanmade
Sponsored
TEA PARTY NEWS
MAKE AMERICA GREAT AGAIN!
JOIN FLORIDA TRUMP RALLIES!
united Muslims of America shared their event
The time has come to understand one simple thing: we the America muslims are as American
Support Hillary. Save American Muslims!
JUL 9
Support Hillary. Save American Muslims!
Sat 1 PM EDT · The Obama White House · Wash...
150 people interested · 47 people going
174 Reactions 36 Comments
Like
Comment
americanmade Florida goes Trump! Here are Florida participating: Daytona Beach, Tampa, Fort Pensacola, Ft ... more

Image showing different posts of social media in the favor of Trump

Google Fact Check Tool



Claim by Social media posts:
"The Left and the Democrats were calling for alternate electors after the 2016 election."

FactCheck.org rating: Misleading
Post Misleadingly Equates 2016 Democratic Effort to Trump's 2020 ...
Jun 29, 2022

Claim by Social media users:
Trump got 74 million votes and there are only 133 million registered voters in the USA

USA Today rating: Partly False

Google Explorer is a cognitive security tool. It is a suite of tools that can be used to detect and mitigate manipulation of search results. However, Google Explorer is not designed to mitigate manipulation post-effect.

Attack Technique 02: Manipulate a Human being.

In this section, we discuss how an attacker manipulates human to get their personal or sensitive data. Manipulation techniques are often used by attackers to deceive and exploit individuals or organizations. It's important to understand these techniques to better protect yourself and your systems.

Bank Scams

Imagine you receive an email that appears to be from your bank. The email claims that there has been suspicious activity on your account and urges you to click on a link to verify your account information and secure your account. The email may use urgent language and a sense of urgency to create a sense of panic and pressure.

However, if you carefully examine the email, you might notice some red flags. The email address of the sender is slightly misspelled or doesn't match the official bank domain. The email may contain grammatical errors or use generic greetings instead of your name. The link provided in the email might lead to a different website that resembles the bank's login page, but it is a clever imitation designed to steal your login credentials.

By clicking on the link and entering your account details, you unknowingly provide your username and password to the attackers. They can then use this information to access your actual bank account, make unauthorized transactions, or steal sensitive information.

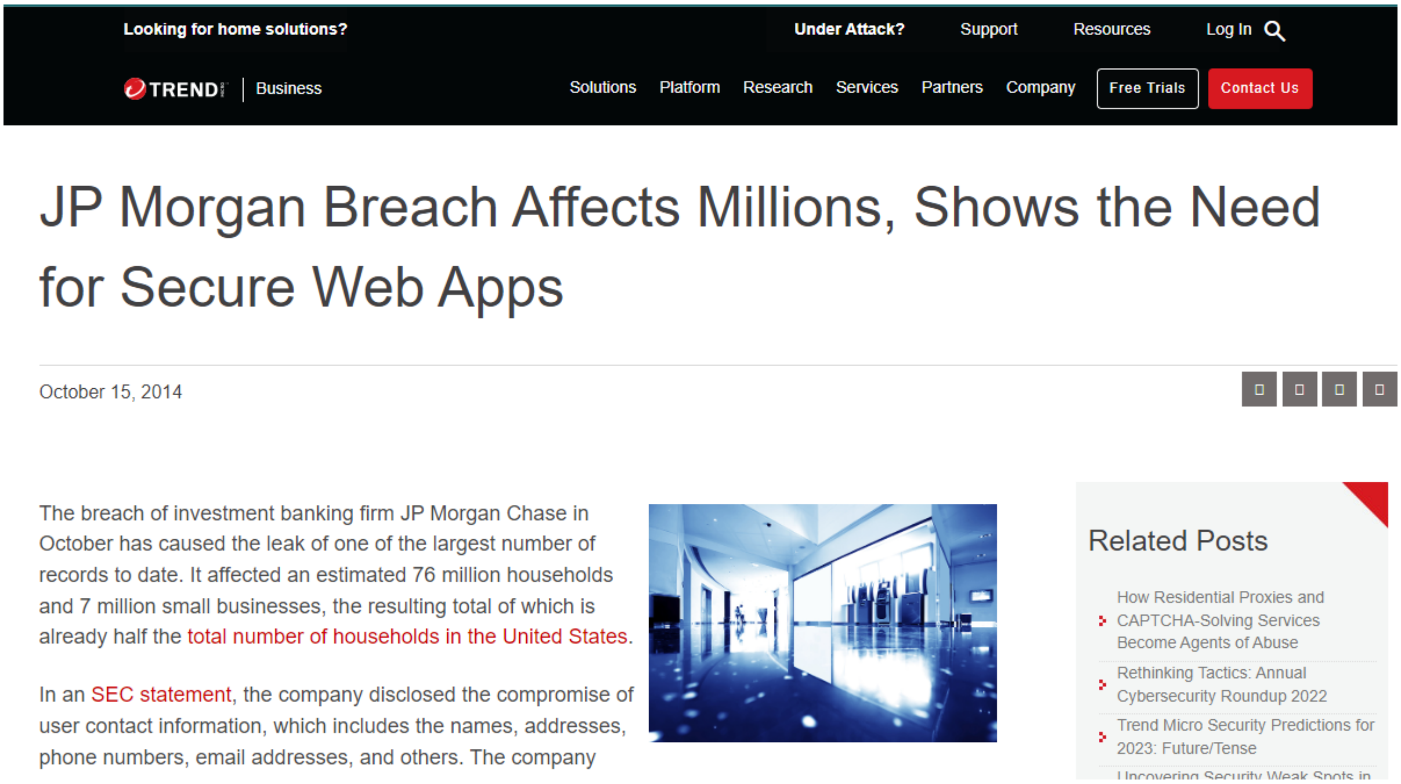
Bank AL Habib
BEWARE
SPAM
Bank AL Habib will never send an email with attachments or links
We will never request for payments or donations
We will never ask your Account Number/PIN/Password/OTP
021-111-014-014 | bankalhabib.com
Bank AL Habib Limited
رشتہ بھروسے کا

This figure shows that Banks will never send you emails or call you for anu PIN/OTP etc.

It's crucial to be cautious and vigilant when dealing with such emails. Always double-check the sender's email address, look for spelling mistakes or grammatical errors, and avoid clicking on suspicious links. When in doubt, contact your bank directly using a trusted contact number or by visiting their official website to verify the authenticity of the message.

Case Study:

In 2014, JPMorgan Chase, a big American bank, experienced a major data breach. Hackers broke into the bank's computer network and stole sensitive information from millions of customers. They got in by tricking employees and using special software to access the network. The stolen data included customer names, addresses, phone numbers, and encrypted passwords. The attack was traced back to a group of hackers from Russia. JPMorgan Chase responded by improving their security and notifying affected customers about the breach. This incident highlighted the importance of strong security measures and quick response to prevent future breaches.



JP Morgan Breach Affects Millions, Shows the Need for Secure Web Apps

October 15, 2014

The breach of investment banking firm JP Morgan Chase in October has caused the leak of one of the largest number of records to date. It affected an estimated 76 million households and 7 million small businesses, the resulting total of which is already half the total number of households in the United States.

In an SEC statement, the company disclosed the compromise of user contact information, which includes the names, addresses, phone numbers, email addresses, and others. The company

Related Posts

- How Residential Proxies and CAPTCHA-Solving Services Become Agents of Abuse
- Rethinking Tactics: Annual Cybersecurity Roundup 2022
- Trend Micro Security Predictions for 2023: Future/Tense
- Uncovering Security Weak Spots in

The image shows a website providing details about the attack on JP Morgan Bank

Conclusion

In conclusion, AMITT is an open-source framework that consists of sets of standards used by both the red team and the blue team to respond to disinformation incidents, and plan strategies to spread, influence, analyze, and detect.

Cognitive security tools are used to detect and analyze disinformation content on different platforms and mitigate the effects of digital harm.

In this section, we try to research and perform a practical that how attackers use different techniques to manipulate human begin and get a positive response to disinformation or sharing their sensitive data. First research shows the tactics that attackers use to spread disinformation for attention and fun. And, tools to detect this disinformation content. In the second practical a phishing call with a colon number and try to manipulate or dismay human beings to share their sensitive information.

Reference

PolitiFact | US, Canadian officials have not released audio of noises heard during Titan search, despite claims

https://github.com/cogsec-collaborative/documentation

https://www.unodc.org/e4j/zh/cybercrime/module-12/key-issues/gender-based-interpersonal-cybercrime.html https://www.ipl.org/essay/JP-Morgan-Chase-Case-Study-PJCWKQZZ2R

https://www.bbc.com/news/technology-46590890

https://www.nytimes.com/2018/12/17/technology/social-media-russia-interference.html

https://www.trendmicro.com/vinfo/us/security/news/cyber-attacks/jp-morgan-breach-affects-millions-shows-need-for-secure-web-apps