How disinformation might hurt your business

And why you are more prepared than you think

Marijn Pronk
is a consultant at KPMG Cyber & Privacy.

Jim Boevink
is a senior consultant at KPMG Cyber & Privacy.

Ellen Mok
is a senior consultant at KPMG Cyber & Privacy.

Disinformation is a cybersecurity threat and should therefore be treated as such. Fortunately, a lot of cybersecurity measures can be adapted to protect your organization from disinformation. This means that we do not need to completely overthrow our already existing cybersecurity strategies. They just need a little tweaking.

INTRODUCTION

Malware, ransomware, Distributed Denial of Service (DDoS), and social engineering are all familiar threats that we associate with cybersecurity. To counter these threats, the cybersecurity industry has come up with many different risk mitigation strategies, frameworks, and tools to protect organizations and businesses.

The bad news is that there is a new cyber threat in town that definitely needs our attention: disinformation. Information being used to deceive the public is not a new concept. However, we can all agree that in this digital age with powerful social media channels, the issue has become exponentially more present in our daily lives.

The good news is that there is no need to completely overthrow all our existing strategies, frameworks, and tools to protect our organization from disinformation attacks. In fact, most of the tools and mitigation measures we use in our organizations to counter the impact of other cyberattacks can also be used to counter the impact of disinformation. They just need a little tweaking.

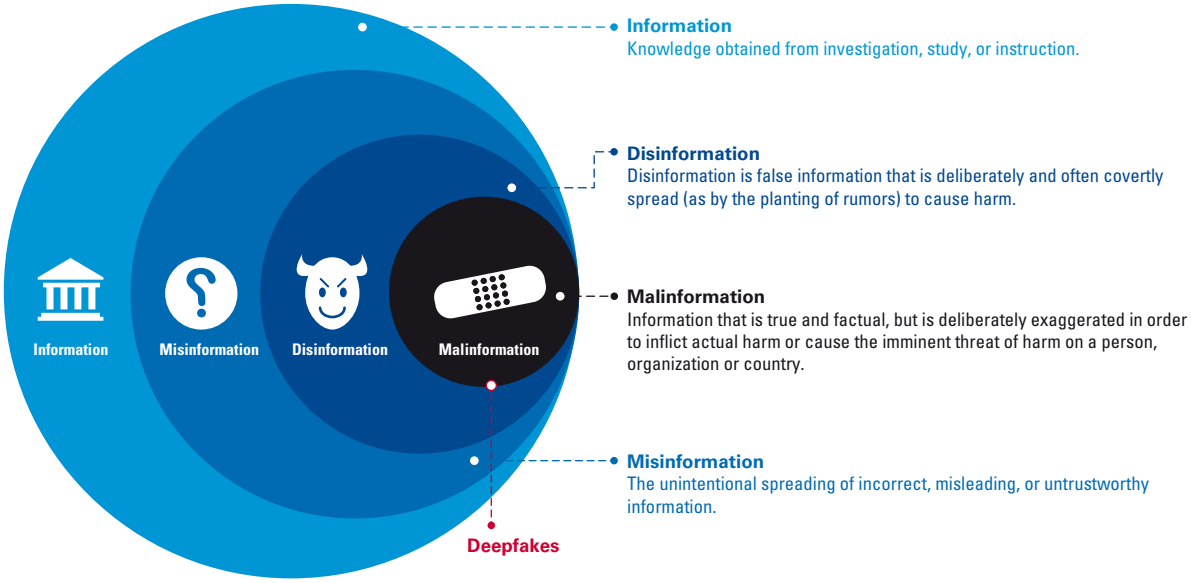
In this article, we will first define disinformation, then we will argue why disinformation should be considered a cyberthreat. Finally, a methodology of preventative measures is proposed to assist your organization in mitigating the effects of disinformation.

TERMINOLOGY OF DISINFORMATION

Before discussing the merits of tweaking your existing cybersecurity strategies and framework regarding this new threat, we first need to establish how to distinguish between the terms misinformation, disinformation, and malinformation. These three concepts together are abbreviated as MDM (Misinformation, Disinformation, Malinformation). In our definition of disinformation, the word "intentional" is key. Whereas misinformation is false information that is spread unintentionally, disinformation is defined as the intentional spreading of misleading or false information with the specific intent to harm or manipulate individuals, groups, or organizations. Finally, we understand malinformation as information that stems from the truth, but is deliberately exaggerated in such a way that it is misleading, and causes harm to a person or an organization. Figure 1 shows the correlation between the different terms.

Disinformation is therefore defined as the intentional spreading of misleading or false information with the specific intent to harm or manipulate individuals, groups, or organizations. However, the process of labeling whether or not a certain piece of information is false or mis-used (misinformation) is mostly irrelevant from a cybersecurity perspective. More relevant in this respect is whether your organization is able to withstand *any kind* of (targeted) information with the (in)direct aim to hurt your business. In doing so, we are keeping away from the ethical and political debate about "freedom of speech versus intentionally spreading false information", and instead focus on how to protect our businesses against this new cyber threat.

Figure 1. Mapping of different information terminologies.



DISINFORMATION AS A CYBER THREAT

A key distinction between a disinformation attack and a "traditional" cyberattack, is the target at which the attack is directed. In most cyberattack methods, attackers use humans to subvert technical systems. In disinformation attacks, technologies are used to subvert human thinking. In fact, it can also be seen as social engineering on a large and automated scale. Some researchers have invented a new word for this: cognitive hacking ([Jiam21]).

Cognitive hacking is a way of using computer or information systems (social media) to change the perceptions and corresponding behaviors of human users ([Bone18]). The attack is aimed at our subconsciousness. In that sense, the goal is the same as with social engineering attacks. There are two major differences, however. Firstly, these cognitive attacks are mostly long-term investments, in that they cause damage that cannot easily be revoked since they are aimed to manipulate the human psyche. Secondly, the tools that are being used are, in part, different.

Threat actors usually deploy three different disinformation tools:

1. social media posts (in some cases pushed by troll farms);
2. fake news sites;
3. deepfakes.

Cognitive hacking is the overarching attack framework in which spreading of disinformation is used as an attack vector.

Bruce Schneier, a fellow and lecturer at Harvard's Kennedy School, wrote two decades ago: "Only amateurs attack machines, professionals target people" ([Schn13]). Phishing and spear phishing are much more effective when a form of cognitive hacking is deployed. Cognitive hacking and other cyberattack methods become more powerful when deployed together. For example, an attacker deploys malware in your IT systems to exfiltrate data and then uses that data to instigate an information campaign against your company with the purpose of extortion or inflicting damage on your brand's reputation and stock value. This exfiltration of data for the purpose of using it to feed a disinformation campaign is especially dangerous for organizations that store substantial amounts of sensitive data (i.e. governmental organizations, health organizations, social media organizations).

The attack chain might also go the other way around: instead of using traditional cyberattacks to exfiltrate data for the purpose of creating disinformation (i.e. spreading disinformation is the end goal), an attacker might use disinformation for a cybersecurity attack (disinformation is used to achieve the goal). For example, an attacker could use false information or false identities (deepfakes) to pretend to be the CEO or CFO, requesting their trusted employees to pay certain invoices that look plausible but actually directly benefit the attacker. CEO/CFO fraud is widely known by now, and different start-up technologies have been developed that can be used to detect deepfakes. That is a good start, but it is also a form of symptom control that only works if organizations approach this risk of disinformation from an overarching information risk management strategy.

HOW TO PROTECT YOURSELF AND YOUR ORGANIZATION AGAINST COGNITIVE HACKING

Although new laws are taking shape ([EC22]), these will not stop perpetrators from using subversive tools for their personal and/or financial gain. That is in

One example of using traditional hacking methods to "feed" a disinformation campaign is the cyberattack on the European Medicines Agency (EMA) of December 2020. The systems of the agency were hacked and confidential internal correspondence on the evaluation of the BioNTech/Pfizer vaccine was unlawfully obtained. Later these documents were published online. However, EMA marked that "Some of the correspondence has been manipulated by the perpetrators prior to publication in a way which could undermine trust in vaccines" ([EMA21]). This shows that traditional cybersecurity attack methods, as in this case hacking, can be used to fuel disinformation campaigns with the aim of carving out trust in public health authorities, such as EMA.

An example where disinformation (deepfakes) is being used in CEO/CFO fraud involves a UK energy firm ([Stup19]). The UK CEO was called, thought that he was speaking to his boss, the chief executive of the parent company. The imitator asked the UK CEO to urgently wire €220,000 to a Hungarian supplier. Afterwards, the company found out that a very realistic deepfake voice (including slight German accent and speech cadence) recording of the chief executive was used by hackers to spoof the CEO. This form of cognitive hacking is not reserved for the most technical criminals, there are many tools online that can convert speech snippets to a workable and believable voice impersonation.

part because, like most of the cyberattack methods, accountability remains a mayor issue, in part due to digital anonymity. This enables users to post and spread disinformation anonymously and it is just as easy to create fake artificial accounts that are able to redirect the disinformation to a much larger audience.

Here is where the previous mentioned tweaking methods come in. In a practical sense, there are some preventive steps that you can take, in the following order:

1. Your organization needs an expansion of your existing information security strategy to include the risk arising from disinformation.
2. Your organization needs to assess the risks and think of scenarios that involve disinformation attacks. These thought exercises can help you understand how your company might be harmed and how prepared your people and processes are to counter these incidents.
3. Your organization needs to determine which mitigating measures are suitable to your organization to counter such risks scenarios, taking into account the nature of your organization's risk profile. You can think of:
   a. creating an incident response plan that sets out how to respond to an MDM attack;
   b. organizing crisis simulations that simulate a large-scale MDM attack;
   c. equipping the PR team and/or your SOC with MDM detection and authentication tools;
   d. training your staff on how to recognize MDM and making them aware of the different ways MDM is deployed.
4. Measure the maturity of your organization to withstand MDM incidents.

Use the structures we already have in place to adapt and protect

CONCLUSION

All of the steps above probably sound very familiar to cybersecurity professionals. That is exactly why we believe that adapting our existing cyber risk management procedures to this emerging threat is achievable for every organization. It all starts with recognizing disinformation as a cyberattack, and not just an elusive political attack deployed by nation-states targeted only at political systems. Cognitive hacking has become a mainstream hacking method and, as such, can be addressed in much the same way as other hacking methods. Instead of trying to reinvent the wheel by implementing specific "disinformation measures", use the structures we already have in place to adapt and protect.

References

[Bone18] Bone, J. (2018, March 5). Cognitive Hack: Trust, Deception and Blind Spots. *Corporate Compliance Insights.* Retrieved from: https://www.corporatecomplianceinsights.com/cognitive-hack-trust-deception-blind-spots/

[EC22] European Commission (2022, November 16). *Digital Services Act: EU's landmark rules for online platforms enter into force.* Retrieved from: https://ec.europa.eu/commission/presscorner/detail/en/IP_22_6906

[EMA21] European Medicines Agency (2021, January 1). *Cyberattack on EMA - update 5.* Retrieved from: https://www.ema.europa.eu/en/news/cyberattack-ema-update-5

[Jiam21] Jiaman, A. (2021, January 30). Disinformation Is a Cybersecurity Threat. *Medium.* Retrieved from: https://medium.com/swlh/disinformation-is-a-cybersecurity-threat-335681b15b48

[Schn13] Schneier, B. (2013, March 1). Phishing Has Gotten *Very* Good. *Schneier On Security.* Retrieved from: https://www.schneier.com/blog/archives/2013/03/phishing_has_go.html

[Stup19] Stupp, C. (2019, August 30). Fraudsters Used AI to Mimic CEO's Voice in Unusual Cybercrime Case. *The Wall Street Journal.* Retrieved from: https://www.wsj.com/articles/fraudsters-use-ai-to-mimic-ceos-voice-in-unusual-cybercrime-case-11567157402

About the authors

Jim Boevink is a senior consultant at KPMG Cyber & Privacy. During his previous employment he worked for the Municipality of Amsterdam where we developed a strategy on how to deal with disinformation aimed at destabilizing local governance. More recently, he has worked on raising awareness on the emerging cyber threat that disinformation is, for organizations and businesses alike, while also working on solutions to mitigate the risk arising from it.

Ellen Mok is a senior consultant at KPMG Cyber & Privacy. She has worked her entire career in the field of cyber security, both in the public and the private sector. Her focus lies on IT Risk Management, cyber security in the public sector, solutions to nation-state security threats and disinformation from a cyber security perspective.

Marijn Pronk is a consultant at KPMG Cyber & Privacy. She has been studying the topic of disinformation for several years and has provided capacity building consultancy directly on the topic of mis-disinformation. She has spoken about disinformation as a cybersecurity threat at the ISF chapter meeting in Antwerp (2022).