

Overview

The production and promulgation of misinformation, disinformation, and malinformation is transforming the public discourse—exacerbating simmering social, political, and geopolitical issues across all sectors of society. These terms are often used interchangeably and erroneously, and threat actors, both government-sponsored and self-serving, are motivated to craft convincing content at dizzying speed, making it difficult to discern fact from fiction.

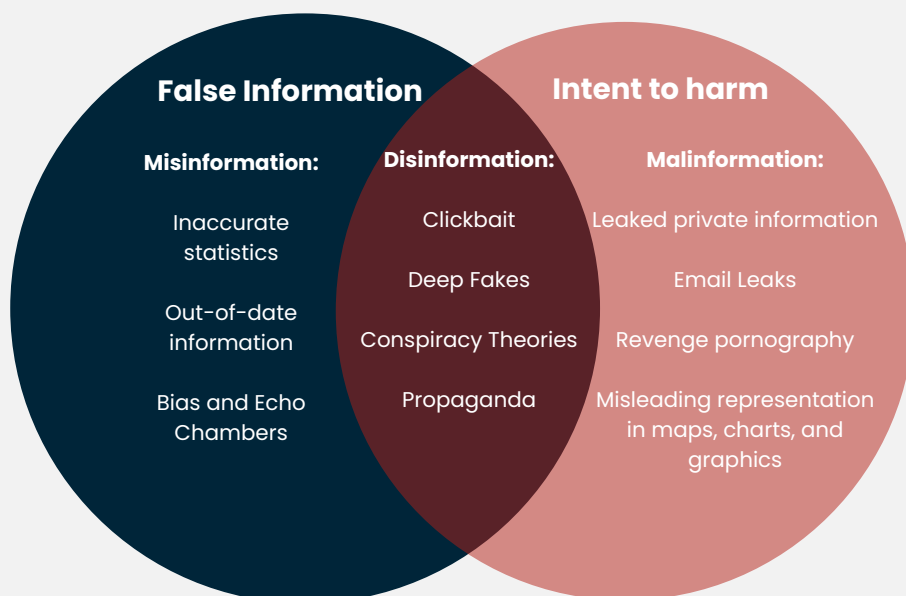
Individuals and organizations must proactively adopt responsible information consumption practices, as well as monitor for sensitive data disclosures online. Tools and resources are available that can assist in verifying the legitimacy of content, and digital risk protection services can assist in identifying and mitigating the spread of sensitive information should such disclosures occur.

Key Findings

- The promulgation of misinformation, disinformation, and malinformation is transforming the public discourse—exacerbating simmering social, political, and geopolitical issues across all sectors of society.
- Misinformation is false information created and shared without the intent to harm or deceive.
- Disinformation is created and promulgated to be deliberately deceptive.
- Malinformation is the disclosure or manipulation of real data and sensitive information that is leaked to cause economic and reputational damage intentionally.

	Definition	Is the information true or based on truth?	Is it spread with intent to cause harm?
Misinformation	Content that is false, misleading, or taken out of context but is shared without the intent to deceive or cause harm	✗	✗
Disinformation	Deliberately false, misleading, manipulated, or biased content spread with the intent to deceive or cause harm	✓✗	✓
Malinformation	Genuine information, or information based on reality, spread with the intent to cause harm	✓	✓

Examples in Practice



In practice, the lines between these categories are not always clear cut.

Disinformation, much like misinformation, often elicits an extreme or reactionary response. The primary difference relates to intent, which can be difficult to determine. While conspiracy theories may be started and circulated for nefarious purposes, they could be propagated by unwitting, misinformed, and well-intentioned individuals that believe it to be true. Conversely, **misinformation** can turn into disinformation when it is still shared by individuals who know it is incorrect but intentionally spread it to stir division.

The boundary between disinformation and **malinformation** is also blurred, particularly when discerning the point at which misrepresentation or distortion of facts becomes outright fabrication.



Misinformation

Content that is false, misleading, or taken out of context but is shared without the intent to deceive or cause harm

Overview

Misinformation is not a new phenomenon, but the internet — and, in particular, social media — has transformed its reach and efficacy. Sharing content quickly and widely is easier now than ever before.

The prevalence of misinformation-online is a product of human and technological factors. Misinformation leverages flaws in human conditioning, a lack of critical challenge, and biases. Users are more likely to react to content that relates to pre-existing grievances and beliefs and are less likely to carefully consider the source — meaning inflammatory subjects generate vast and rapid engagement.

Technological factors amplify misinformation online. Social media and news feeds rely on deep learning and statistical algorithms, such as recommender systems, that can enable similarly-behaving users to be grouped and content to be promoted that they are more likely to react to.^[1] This includes content with high prior engagement and that is likely to elicit strong emotional reactions. A post that receives strong initial response will be shown to more users, who then also demonstrate strong interaction — amplifying online misinformation. This is also the case with disinformation.

Consequences

Despite well-meaning intentions, misinformation has negative consequences. It prevents users from making informed decisions, often resulting in people acting against their own best interest. It can drive an unhealthy skepticism towards all information received; excessive distrust of politicians, journalists, and experts; and the erosion of democratic processes.

Themes

Misinformation pertains to most emotive topics. Popular themes include armed conflicts, politics, celebrities, sports, and businesses.



Examples

The COVID-19 pandemic created a perfect storm of online misinformation, with evolving scientific consensus and an anxious population. For example, following new studies that proved mask efficacy against spreading the disease, the U.S. Centers for Disease Control and Prevention updated guidance on mask use. Users continued to propagate out-of-date guidance based on old information that affirmed their belief masks were ineffective.^[2, 3]

Shortly after Russia's invasion of Ukraine, social media users misrepresented footage from a video game showing close combat between fighter jets, claiming the clip showed the Ukrainian air force fighter pilot dubbed the "Ghost of Kyiv" shooting down a Russian air force fighter jet.^[4] It is not clear if this began with the intent to deceive but was propagated by many individuals that likely thought the footage was legitimate. The video was flagged as "out of context" by Twitter, although similar content on Ukraine's official Twitter account did not get flagged.

Satirical websites and programs are frequently misinterpreted and shared as fact. The Onion is so regularly misunderstood and shared on social media as presenting legitimate content that an online community has been created to mock those who were fooled.^[5]



Footage incorrectly alleged to be the "Ghost of Kyiv", engaged with by over 20,000 users and flagged as "out of context" by Twitter

Source:
hXXps://twitter.com/ItsOlegi21/status/1497164782952239104?s=20&t=NfaX2G2awdaWhkEKAQ2O_g

Key Identifiers

- Headlines that seem shocking or outrageous, which may be attempting to exploit viewers' emotions.
- Content with mistakes or that looks doctored.
- Content that appears funny or sarcastic.

Recommendations to Combat

- Question the credibility of the source.
- Consider the date the content was published.
- Consider the content in its entirety, not just a headline.
- Use fact-checking websites, like FactCheck.org, the News Literacy Project, and NewsGuard.
- Conduct training on how misinformation works to boost defenses to, and recognition of, false claims.
- For organizations that find themselves a victim of misinformation, have a crisis response plan in place to evaluate the content and formulate an approach to neutralize and contain the misinformation.

FOOTNOTES

- [1] hXXp://ceur-ws.org/Vol-2758/OHARS-paper3.pdf
- [2] hXXps://www.factcheck.org/2020/04/covid-19-face-mask-advice-explained/
- [3] hXXps://www.factcheck.org/2021/03/scicheck-the-evolving-science-of-face-masks-and-covid-19/
- [4] hXXps://apnews.com/article/fact-checking-548180451272?utm_campaign=SocialFlow&utm_source=Twitter&utm_medium=APFactCheck
- [5] hXXps://www.reddit.com/r/AtTheOnion/



Disinformation

Deliberately false, misleading, manipulated, or biased content spread with the intent to deceive or cause harm

Overview

Disinformation is created to be deliberately deceptive, and content that is labelled as “fake news” often falls within this category – at least, that is the assertion being made by the labeller. Disinformation campaigns are typically motivated by three factors: power or influence, profit, or to sow division. Campaigns are widespread and conducted by actors ranging from nation states to online trolls. Disinformation can run afoul of restrictions on free speech – such as slander, libel, and hate speech, which may not be protected under the U.S. First Amendment.

Threat actors can leverage disinformation in social engineering scams. These scams can serve as an intrusion vector for a plethora of malicious activity, including stealing user credentials, deploying payloads, and conducting payment fraud. Impersonations and fake reviews are also considered disinformation. Cyber threat actors may claim to have conducted attacks that never took place to increase their reputation or to elicit payment from victims, regardless.

Consequences

Disinformation is designed to have negative consequences. The discreditation of legitimate, impartial outlets pushes people towards alternative sources that may lack credibility. Campaigns can also result in financial and reputational loss, both at an individual and enterprise level, and lead to additional compromise.

Case Study: Fact-checking the “Fact Checkers”

Fact-checking websites and sources have been established to combat the spread of disinformation online. However, there are multiple examples where these have been set up to appear credible and hide an obvious and detrimental bias. During a televised election debate in 2019, UK Prime Minister Boris Johnson’s Political Party renamed the @CCHQPress account – an account verified by Twitter and controlled by the Conservative Party – to appear as a fact-checking service. This is a clear attempt to deceive an audience by not only pretending to be a different entity but also deliberately obfuscating the biased nature of the account.



Themes

Disinformation can be found online relating to most emotive topics. Popular themes include military conflicts, politics, celebrities, sports, and businesses.

Examples

Nation-states frequently leverage disinformation campaigns, particularly to exert power or influence outcomes. As early as 2018, a dozen Russian military intelligence officers were indicted for their alleged roles in interfering with the 2016 U.S. elections.^[6] Reporting stated that Russia leveraged bots to disseminate more than 3.8 million tweets – about 19 percent of the total tweets related to the 2016 U.S. presidential election – with ~80 percent supporting Donald Trump.^[7]

In June 2022, a Chinese threat actor dubbed DragonBridge was identified unsuccessfully attempting to mobilize public protests against an Australian mining company – very likely in an attempt to defend China’s dominance in the market.^[8] The group has been linked to past influence campaigns, including amplifying claims made by Russian influence operations that the United States was operating biological weapons labs in Ukraine.

In August 2022, Meta and Twitter removed accounts engaged in disinformation that promoted pro-Western narratives to social media users in the Middle East and Central Asia.^[9]

Disinformation has been widespread during the war in Ukraine. One of the main premises for Russia’s invasion was the allegation that Ukraine was carrying out genocide in Donbas – claims that have since been debunked as baseless. In March 2022, Meta removed a deepfake of Ukrainian President Zelensky in which he appeared to call on the military to lay down their arms.^[10] Ukraine has also spread disinformation, such as the Snake Island stand – likely an attempt to boost morale. In August 2022, Ukraine shut down a bot farm of 1,000,000 bots used to spread disinformation on social networks.^[11]



Deep fake of President Zelensky, March 2022
Source:
[hXXps://twitter.com/MikaelThalen/status/1504123674516885507](https://twitter.com/MikaelThalen/status/1504123674516885507)

In May 2022, researchers identified that Fronton, a Russian Internet-of-Things botnet designed to conduct distributed denial-of-service (DDoS) attacks, could be used to conduct mass-scale social media disinformation campaigns.^[12] Fronton leverages a web-based dashboard that is designed to create fraudulent social media profiles and manufacture fake breaking news.

FOOTNOTES

- [6] [hXXps://www.justice.gov/opa/pr/grand-jury-indicts-12-russian-intelligence-officers-hacking-offenses-related-2016-election](https://www.justice.gov/opa/pr/grand-jury-indicts-12-russian-intelligence-officers-hacking-offenses-related-2016-election)
- [7] [hXXps://firstmonday.org/article/view/7090/5653](https://firstmonday.org/article/view/7090/5653)
- [8] [hXXps://thehackernews.com/2022/07/pro-china-group-uses-dragonbridge.html](https://thehackernews.com/2022/07/pro-china-group-uses-dragonbridge.html)
- [9] [hXXps://www.bloomberg.com/news/articles/2022-08-24/meta-removes-fake-accounts-tied-to-pro-us-influence-campaign#xj4y7vzk](https://www.bloomberg.com/news/articles/2022-08-24/meta-removes-fake-accounts-tied-to-pro-us-influence-campaign#xj4y7vzk)
- [10] [hXXps://www.infosecurity-magazine.com/news/russia-uses-deepfake-zelensky/](https://www.infosecurity-magazine.com/news/russia-uses-deepfake-zelensky/)
- [11] [hXXps://www.bleepingcomputer.com/news/security/ukraine-takes-down-1-000-000-bots-used-for-disinformation/](https://www.bleepingcomputer.com/news/security/ukraine-takes-down-1-000-000-bots-used-for-disinformation/)
- [12] [hXXps://thehackernews.com/2022/05/fronton-russian-iot-botnet-designed-to.html](https://thehackernews.com/2022/05/fronton-russian-iot-botnet-designed-to.html)



Threat actors are able to leverage disinformation scams for monetary gain. This includes leveraging fake or misleading information to solicit donations to a fake charity, as has been seen in multiple cases following Russia's invasion of Ukraine.^[13] In one recently detailed case, threat actors that were likely based in Vietnam established malicious news sites to peddle QAnon conspiracy theories and U.S. political disinformation. The operators generated revenue via advertising, affiliate sales links, and direct solicitation of donations. Despite having a financial motive, campaigns of this nature create increased political division and amplify conspiracy theories and disinformation.^[14]

Financially-motivated threat actors are also increasingly leveraging disinformation to conduct cryptocurrency scams, with more than 46,000 people reporting losses of over USD one billion to scams and an average median loss of USD 2,600.^[15] In May 2022, threat actors leveraged deep fake videos of Elon Musk and other prominent cryptocurrency players to promote a scam involving BitVex trading platform that enabled theft of deposited currency.^[16]



Deep fake video of Elon Musk promoting BitVex trading platform

Source:

[hXXps://heimdalsecurity\[.\]com/blog/deep-fakes-of-elon-musk-promote-bitvex-fraud/](https://heimdalsecurity.com/blog/deep-fakes-of-elon-musk-promote-bitvex-fraud/)

Malinformation

Genuine information, or information based on reality, spread with the intent to cause harm

Overview

Malinformation poses a significant threat to the public, VIPs, and organizations in almost all industries. Like disinformation, malinformation actors are typically driven by power or influence, profit, or to sow division. It differs from disinformation in that it contains true or genuine information.

Although end users may have been desensitized to malinformation, the potential impact from leaked information is on an upward trajectory. This is driven by greater interconnectivity of online networks and services. Threat actors leverage multiple means of stealing data, ranging from infostealers and trojans to ransomware.

Attacks beget further attacks. Owing to the near-ubiquitous use of double extortion tactics, ransomware has become one of the main drivers of data breaches. Private data such as private emails can be leveraged by threat actors to extort payments from victims.

Consequences

Malinformation can lead to reputational damage and financial loss, both at an individual and enterprise level. It can be used to undermine political and corporate negotiations, influence democratic processes, and compromise national security.

Themes

Examples of malinformation include, but are not limited to, dissemination of confidential information, leaking private emails, release of stolen intellectual property, and revenge porn.

Examples

Malinformation can be leveraged by nation-state actors as an instrument of war. In March 2022, the intelligence arm of the Ministry of Defense of Ukraine (GURMO) hacked and leaked documents it claimed referred to intellectual property concerning the Russian Belayarsk Nuclear Power Station.^[17]

The 2016 Panama Papers disclosures are classified as malinformation, despite the whistleblowing-intentions of the journalists.^[18] The papers were stolen from law firm Mossack Fonseca and outline sensitive financial and legal information for offshore entities.

In 2017, French President Emmanuel Macron's campaign was hacked days before the French presidential election, with more than 9GB of data dumped on a file-sharing site.^[19] The leaked emails claimed to show evidence of criminal wrongdoing, including tax evasion and election fraud, though were later widely assessed to be mundane.

Key Identifiers

- Headlines that seem shocking or outrageous may be attempting to exploit viewers' emotions.
- Content with mistakes or that looks doctored.
- Content that appears funny or sarcastic.

Recommendations to Combat

- Question the credibility of the source.
- Consider the date the content was published.
- Consider the content in its entirety, not just a headline.
- Use fact-checking websites, like FactCheck.org, the News Literacy Project, and NewsGuard.
- Conduct training on how disinformation works to boost defenses to, and recognition of, false claims.
- For organizations that find themselves a victim of disinformation, have a crisis response plan in place to evaluate the content and formulate an approach to neutralize and contain the disinformation.

FOOTNOTES

^[13] [hXXps://www.bbc\[.\]co.uk/news/uk-61078742](https://www.bbc.com/news/uk-61078742)

^[14] [hXXps://www.isdglobal.org/wp-content/uploads/2022/01/Conspiracy-Clickbait-Study-3.pdf](https://www.isdglobal.org/wp-content/uploads/2022/01/Conspiracy-Clickbait-Study-3.pdf)

^[15] [hXXps://www.ftc.gov/news-events/data-visualizations/data-spotlight/2022/06/reports-show-scammers-cashing-crypto-craze](https://www.ftc.gov/news-events/data-visualizations/data-spotlight/2022/06/reports-show-scammers-cashing-crypto-craze)

^[16] [hXXps://heimdalsecurity\[.\]com/blog/deep-fakes-of-elon-musk-promote-bitvex-fraud/](https://heimdalsecurity.com/blog/deep-fakes-of-elon-musk-promote-bitvex-fraud/)

^[17] [hXXps://www.somagazine\[.\]com/analysis/breach/in-a-first-ukraine-leaks-russian-intellectual-property-as-act-of-war](https://www.somagazine.com/analysis/breach/in-a-first-ukraine-leaks-russian-intellectual-property-as-act-of-war)

^[18] [hXXps://www.icij.org/investigations/panama-papers/](https://www.icij.org/investigations/panama-papers/)

^[19] [hXXps://www.reuters\[.\]com/article/us-france-election-macron-leaks-idUSKBN1812AZ](https://www.reuters.com/article/us-france-election-macron-leaks-idUSKBN1812AZ)



During the 2018 Brazilian Presidential election, Jair Bolsonaro used malinformation to discredit his main political rival, Fernando Haddad. Bolsonaro's camp propagated the fabrication that Haddad, a former Minister for Education then running for President, had created a "gay kit" to be introduced in elementary schools to encourage children to become gay.^[20] The malinformation came from a distortion of Haddad's promotion of an educational programme, while Minister for Education, for elementary school students to be taught about sexual diversity and to tackle homophobia.

In 2021, a news outlet edited a video of U.S. President Joe Biden to remove context and make it seem that he was making racially insensitive remarks, before claiming Biden was "facing backlash" about his comments.^[21] The hashtag #RacistJoeBiden was trending on Twitter shortly after. Despite claiming the edit was made due to time constraints, it is more likely the malinformation was designed to disseminate the false narrative that Biden was revealing a negative bias against Black Americans and sow political discord.



News' segment including the edited version President Biden's address that removed context to his comments

Source: [hXXps://twitter.com/atruper/status/1459198578358956037](https://twitter.com/atruper/status/1459198578358956037)

Key Identifiers

- Sensitive information appears on Dark Web marketplaces and leak sites.
- Sensitive data is shared in open sources, such as social media.

Recommendations to Combat

- Ensure a thorough understanding of the exposure of key individuals and organizations in open and closed sources.
- Keep up to date with evolving threat actor tactics, techniques, and procedures.
- Proactively monitor for potential network access sales to obtain critical early warnings of impending cyberattacks.
- Monitor news and social media for breaking stories on data breaches and leaks.
- Employ digital risk protection services to determine the credibility of breach claims and mitigate potential impacts.
- For organizations that find themselves a victim of malinformation, have a crisis response plan in place to evaluate the content and formulate an approach to neutralize and contain the malinformation.

Appendix: Traffic Light Protocol for Dissemination

TLP:RED

HOW IT IS USED

Sources may use **TLP:RED** when information cannot be effectively acted upon by additional parties and could lead to impacts on a party's privacy, reputation, or operations if misused.

HOW IT IS SHARED

Recipients may **NOT** share **TLP:RED** with any parties outside of the specific exchange, meeting, or conversation in which it is originally disclosed.

TLP:AMBER

HOW IT IS USED

Sources may use **TLP:AMBER** when information requires support to be effectively acted upon but carries risks to privacy, reputation, or operations if shared outside of the organizations involved.

HOW IT IS SHARED

Recipients may **ONLY** share **TLP:AMBER** information with members of their own organization and its clients, but only on a need-to-know basis to protect their organization and its clients and prevent further harm. **Note that TLP:AMBER+STRICT** restricts sharing to the organization only.

TLP:GREEN

HOW IT IS USED

Sources may use **TLP:GREEN** when information is useful for the awareness of all participating organizations, as well as with peers within the broader community or sector.

HOW IT IS SHARED

Recipients may **ONLY** share **TLP:GREEN** information with peers and partner organizations within their sector or community but not via publicly accessible channels.

TLP:CLEAR

HOW IT IS USED

Sources may use **TLP:CLEAR** when information carries minimal or no risk of misuse in accordance with applicable rules and procedures for public release.

HOW IT IS SHARED

Recipients may share **TLP:CLEAR** information without restriction, subject to copyright controls.

ZEROFOX INTELLIGENCE
September 9, 2022

FOOTNOTES

- [20] [hXXps://www.sbs.com.au/news/dateline/article/how-jair-bolsonaro-used-fake-news-to-win-power/c0upl3n5k](https://www.sbs.com.au/news/dateline/article/how-jair-bolsonaro-used-fake-news-to-win-power/c0upl3n5k)
[21] [hXXps://www.theguardian.com/media/2021/nov/13/fox-news-edits-biden-video-negro-leagues-satchel-paige](https://www.theguardian.com/media/2021/nov/13/fox-news-edits-biden-video-negro-leagues-satchel-paige)

