

Icebreaker

**Where do you or your
partners and target
groups get news?**

<https://padlet.com/swapneelsmehta/where-do-you-and-your-partners-target-groups-get-your-news-f-s2lo2ggbk-cuj57hx>

Disinformation: Threats and Mitigation

Friday 24/03/2023

Swapneel Mehta

One Fact Foundation

“Challenges and Opportunities for Media Development”



About Me



2019-23 Data Science Ph.D. Candidate @ New York University

2022 Early Detection of Misinfo. @ Twitter Civic Integrity

2021 Trending Hashtag Recommendation @ Adobe

2018-19 Machine Learning x Particle Physics @ CERN



@swapneel_mehta



Mis, Dis and Malinformation

MDM Archetypes

Misinformation

Inaccurate or misleading
information (no intent)



Disinformation

Misinformation +
intent to deceive



Malinformation

True information +
intent to deceive



Mis, Dis and Malinformation

MDM Archetypes

Misinformation ————■

Inaccurate or misleading
information (no intent)

■ ———— **Disinformation**

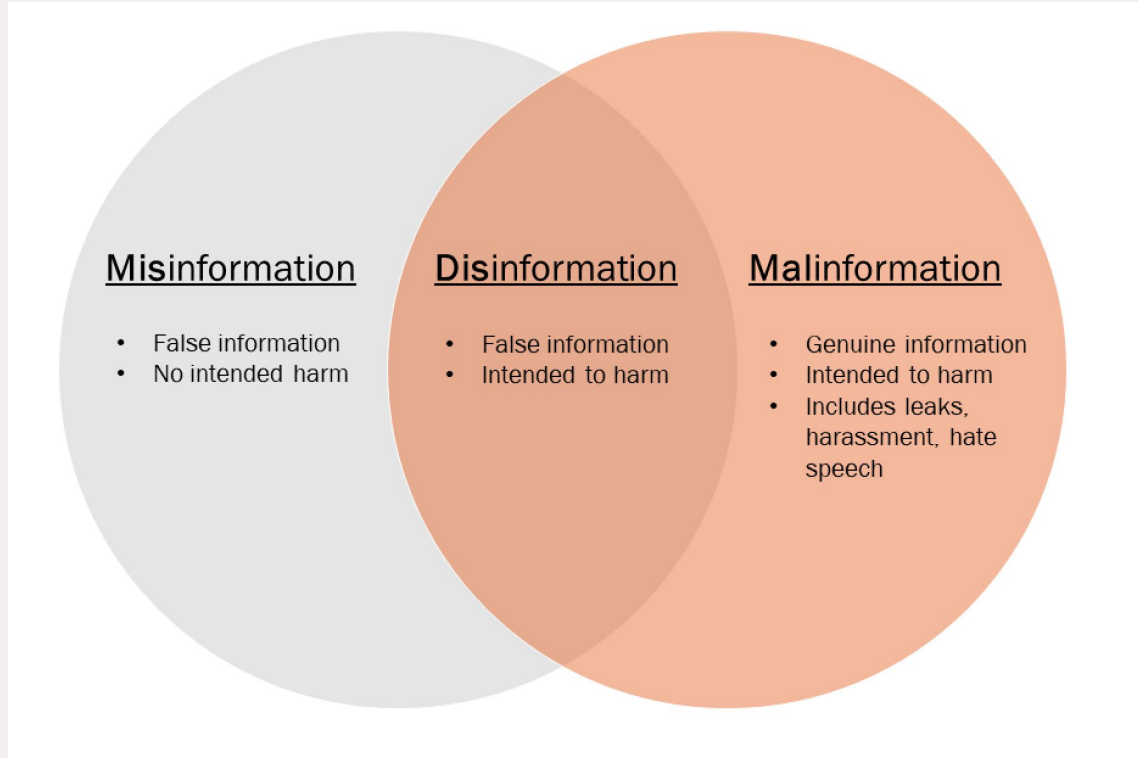
Misinformation +
intent to deceive

Malinformation ————■

True information +
intent to deceive

Mis, Dis and Malinformation

MDM Archetypes



Source: Council of Europe's Information Disorder framework.

What would you classify this as?

MDM Examples

The screenshot shows the Newsmax website interface. At the top, there are navigation links for NEWSMAX, TV, HEALTH, FINANCE, and WORLD. The main Newsmax logo is prominently displayed in the center. Below the logo, the date "Monday, March 20, 2023" is shown. A horizontal menu contains links for Home, Platinum, Newsfront (which is highlighted), Politics, Opinion, Podcast, The Wire, Books, Best Lists, Specials, and Sci & Tech. Below the menu, the breadcrumb "Home | Newsfront" is visible. The main article headline reads "SVB Donated \$73M to 'BLM Movement'", with tags for "svb", "black lives matter", "esg", and "social justice". On the right side, there is a "Free Newsmax E-Alerts" sign-up form with fields for Email, Country (set to United States), and Zip Code, along with a "SIGN UP" button and a privacy notice: "Privacy: We never share your email."


<https://app.sli.do/event/1hXNLSYVS2d9Uy8vGeuB7N/embed/polls/d14a6f68-6a06-49b9-9662-a4b07d047215>

What would you classify this as?

MDM Examples

Newsmax (News)

AllSides Media Bias Rating™: **Right** ✓ agree ✗ disagree



4
Right

LLCRR

What does this mean?

How we determined this rating:

- ✓ **Independent Review**
- ✓ **Editorial Review:** Jan 2021, Dec 2019
- ✓ **Community Feedback:** 24,397 ratings


■ AllSides has **high** confidence in this bias rating.

Unless otherwise noted, this bias rating refers only to online news coverage, not TV, print, or radio content.

[Learn about our bias rating methods →](#)


Contents

- [1. About Newsmax \(News\)](#)
- [2. What a "Right" Rating Means](#)
- [3. Details](#)
- [4. Bias Reviews](#)
- [5. Community Feedback](#)
- [6. Confidence Level](#)
- [7. Additional Information](#)
- [8. Ownership & Funding](#)
- [9. Articles from Newsmax \(News\)](#)



Bias Rating	Right ⓘ
Type	News Media
Owner	Newsmax Media
Established	1998
Website	newsmax.com
Twitter	@newsmax
Facebook	newsmax
Wikipedia	Newsmax (News)

[Go to Newsmax \(News\) ↗](#)

 Possible Paywall

What would you classify this as?

MDM Examples



A screenshot of a tweet from Donald J. Trump. The tweet text reads: "Heartwarming to see all of the tremendous support out there, especially the organic Rallies that are springing up all over the Country, including a big one on Saturday in D.C. I may even try to stop by and say hello. This Election was Rigged, from Dominion all the way up & down!". Below the text is a blue warning icon and the text "This claim about election fraud is disputed".

Donald J. Trump ✓
@realDonaldTrump

Heartwarming to see all of the tremendous support out there, especially the organic Rallies that are springing up all over the Country, including a big one on Saturday in D.C. I may even try to stop by and say hello. This Election was Rigged, from Dominion all the way up & down!

⚠ This claim about election fraud is disputed

<https://app.sli.do/event/1hXNLSYVS2d9Uy8vGeuB7N/embed/polls/d14a6f68-6a06-49b9-9662-a4b07d047215>

What would you classify this as?



[Let's watch the Video](#)

<https://app.sli.do/event/1hXNLSYVS2d9Uy8vGeuB7N/embed/polls/49e7e3e5-0de6-4841-ba9a-e90c019a5252>

What would you classify this as?

MDM Examples



The screenshot shows a social media post from Mike Lamb (@ML73050761). The text of the post reads: "Joe Biden falls down Air Force One stairs, falls up Air Force One Stairs, gets lost on red carpets. Weeble Wobbles wobble but don't fall down like Joe Biden." Below the text is a video player showing a scene from an airport tarmac at night. A large blue and white Air Force One aircraft is the central focus. The video player interface includes a play button, a progress bar showing 0:01 / 1:08, and a volume icon. Below the video player, there is a yellow banner with the text "PREZYDENT USA W POLSCE" and "SAMOLOT BIDENA NA LOTNISKU CHOPINA W WARSZAWIE". The video has 146 views and was posted at 11:58 PM on Feb 22, 2023, with 284 views.

[Let's watch the Video](#)

<https://app.sli.do/event/1hXNLSYVS2d9Uy8vGeuB7N/embed/polls/49e7e3e5-0de6-4841-ba9a-e90c019a5252>

Modalities of MDM Operations



Text



Audio



Images



Video

Thinking through a Historical Lens



Text

Newspapers



Image

Paintings



Audio

Radio



Video

Television

Modalities of Disinformation



1999-07

Texting



2003-08

Image-sharing



2000-12

Podcasts



2005-10

Video

Discussion

**Which modality is most effective to spread
disinformation?**

Discussion

JOURNAL ARTICLE

Seeing Is Believing: Is Video Modality More Powerful in Spreading Fake News via Online Messaging Apps?

S Shyam Sundar , Maria D Molina, Eugene Cho

Takeaway

- **Video is more persuasive than text and audio:**
 - Text needs people to read, visualize, imagine, and understand
 - Video is upfront, ‘see to believe’
 - Video is associated with greater engagement and higher perceived realism

Discussion

Why do people share fake news?

<https://padlet.com/swapneelsmehta/why-do-people-share-fake-news-t5oe2j46ph44af8w>

Discussion

The psychological drivers of misinformation belief and its resistance to correction

[Ullrich K. H. Ecker](#) , [Stephan Lewandowsky](#), [John Cook](#), [Philipp Schmid](#), [Lisa K. Fazio](#), [Nadia Brashier](#), [Panayiota Kendeou](#), [Emily K. Vraga](#) & [Michelle A. Amazeen](#)

Takeaway

Cognitive drivers

Intuitive thinking

- Lack of analytical thinking and/or deliberation

Cognitive failures

- Neglect source cues and/or knowledge
- Forget source and/or counter-evidence

Illusory truth

- Familiarity
- Fluency
- Cohesion

False beliefs

Socio-affective drivers

Source cues

- Elites
- In-group
- Attractive

Emotion

- Emotive information
- Emotional state

Worldview

- Personal views
- Partisanship

Methods of Proliferation of MDM Ops

Adversarial Actors

Agenda	Approach
Dismiss	Discrediting true information that contradicts disinfo.
Distort	Taking true info. and altering or exaggerating it.
Distract	Deliberately lying / creating false info. to deceive others.
Dismay	Using scare tactics to prevent the truth from coming out
Divide	Employ polarization tactics to exploit divisive opinions

Manifestation of MDM Ops

Adversarial Behaviors

Behavior	Description
SockPuppet	Online identity used to deceive, typically in public
Astroturfing	Taking true info. and altering or exaggerating it.
Sealioning	Pester people with persistent requests for evidence (JAQ)
Catfishing	Online identity used to deceive, targeting individuals

Manifestation of MDM Ops

Table 1 Communication objectives BEND


	Manipulating the narrative		Manipulating the social network	
Positive	Engage	Messages that bring up a related but relevant topic	Back	Actions that increase the importance of the opinion leader or create a new opinion leader
	Explain	Messages that provides details on or elaborate the topic	Build	Actions that create a group or the appearance of a group
	Excite	messages that elicit a positive emotion such as joy or excitement	Bridge	Actions that build a connection between two or more groups
	Enhance	Messages that encourage the topic-group to continue with the topic	Boost	Actions that grow the size of the group or make it appear that it has grown
Negative	Dismiss	Messages about why the topic is not important	Neutralize	Actions decrease the importance of the opinion leader
	Distort	Messages that alter the main message of the topic	Nuke	Actions that lead to a group being dismantled or breaking up, or appearing to be broken up
	Dismay	Messages that elicit a negative emotion such as sadness or anger	Narrow	Actions that lead to a group becoming sequestered from other groups or marginalized
	Distract	Discussion about a totally different topic and irrelevant	Neglect	Actions that reduce the size of the group or make it appear that the group has grown smaller

Countering MDM Operations

So how can we stop these tactics?

Discussion

- **Shifting attention to accuracy can reduce misinformation online**

[Gordon Pennycook](#) , [Ziv Epstein](#), [Mohsen Mosleh](#), [Antonio A. Arechar](#), [Dean Eckles](#) & [David G. Rand](#) 

Takeaway

- **People say sharing only accurate information is important**
 - Sharing does not necessarily reflect belief
 - Unique intervention – follow people and prompt them to identify accuracy of headline
 - Improves the quality of shared information after the intervention
 - Do people just need nudges?

Identification and Mitigation

Category	Approach	Examples
Fact-checking	Manually verify claims in posts	<u>IFCN</u> , <u>Open Feedback</u> , <u>Politifact</u> , <u>Snopes</u> , <u>FactCheck.org</u> ,
Scores	Scoring the credibility of media providers or accounts	<u>MBFC</u> , <u>Iffy News</u> , <u>NewsGuard</u> , <u>Botometer</u> , <u>Have I shared Fake News</u>
Policy Interventions	Platform content moderation, external	<u>Twitter</u> , <u>Meta</u> , <u>Reddit</u> , <u>Media Literacy</u> (Interventions)
External datasets and dashboard audits	Build bespoke games, dashboards for civic datasets	<u>Bellingcat Ukraine</u> , <u>Hoaxy</u> , <u>TBCOV</u> , <u>Bad News Game</u> , <u>Inoculation Game</u>
Coordinated Inauthentic Behavior	Detecting Networks of superspreaders and bots	<u>IO Archive</u> , <u>ASPI</u> , <u>Transparency Reports</u>

Frameworks to Think Through

<https://disarmframework.herokuapp.com>