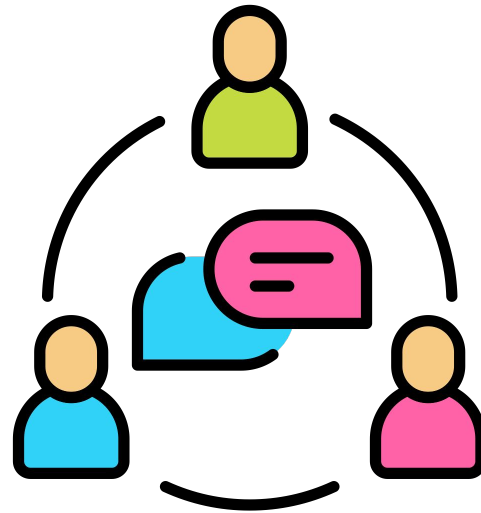# 1.1

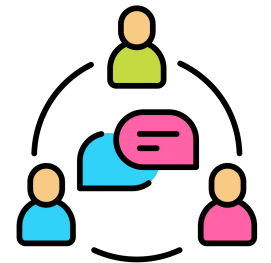# Disinformation, Misinformation, Malinformation and Fake News

Note

- Interactive symbol: tasks for groups

Online Harms according to the World Economic Forum

- Build groups: How would you define the term "Online Harms"? Which online phenomena would fall under your definition? Make a short list

12/05/2025

Online Harms according to the World Economic Forum

- **Definition**
  - **Threats to personal and community safety**, harms to health and well-being, hate and discrimination, violation of dignity, invasion of privacy, deception and manipulation (World Economic Forum, 2023)
  - E. g., grooming, pro-terror material, extremism recruitment, radicalisation, violence, dangerous behaviours, technology-facilitated abuse, gender-based violence

12/05/2025

Online Harms according to the World Economic Forum

- **Definition**
  - Threats to personal and community safety, **harms to health and well-being**, hate and discrimination, violation of dignity, invasion of privacy, deception and manipulation (World Economic Forum, 2023)
  - E. g., suicide, self-injury, eating disorders, developmentally inappropriate content

12/05/2025

Online Harms according to the World Economic Forum

- **Definition**
  - Threats to personal and community safety, harms to health and well-being, **hate and discrimination**, violation of dignity, invasion of privacy, deception and manipulation (World Economic Forum, 2023)
  - E. g., hate speech, algorithmic discrimination
  - Hate speech: Using derogatory or discriminatory language when referencing to an individual or a group of people, especially attacking their protected identity characteristics, including but not limited to gender, religion, and ethnicity

12/05/2025

Online Harms according to the World Economic Forum

- **Definition**
  - Threats to personal and community safety, harms to health and well-being, hate and discrimination, **violation of dignity**, invasion of privacy, deception and manipulation (World Economic Forum, 2023)
  - E. g., online bullying, harassment, sexual extortion

12/05/2025

Online Harms according to the World Economic Forum

- **Definition**
  - Threats to personal and community safety, harms to health and well-being, hate and discrimination, violation of dignity, **invasion of privacy**, deception and manipulation (World Economic Forum, 2023)
  - E. g., doxxing, Image-based abuse
  - Doxxing: Exposing sensitive identity details online without the victim's consent with the intention of placing the target at risk, intimidating, and doing harm

12/05/2025

Online Harms according to the World Economic Forum

- **Definition**
  - Threats to personal and community safety, harms to health and well-being, hate and discrimination, violation of dignity, invasion of privacy, **deception and manipulation** (World Economic Forum, 2023)
  - E. g., Disinformation, misinformation, "deceptive synthetic media", scams, impersonation, phishing, catfishing
  - Deepfakes: Content, esp. videos, generated to appear real with the help of artificial intelligence and analysis of large amounts of data (Armitage & Vaccari, 2021)
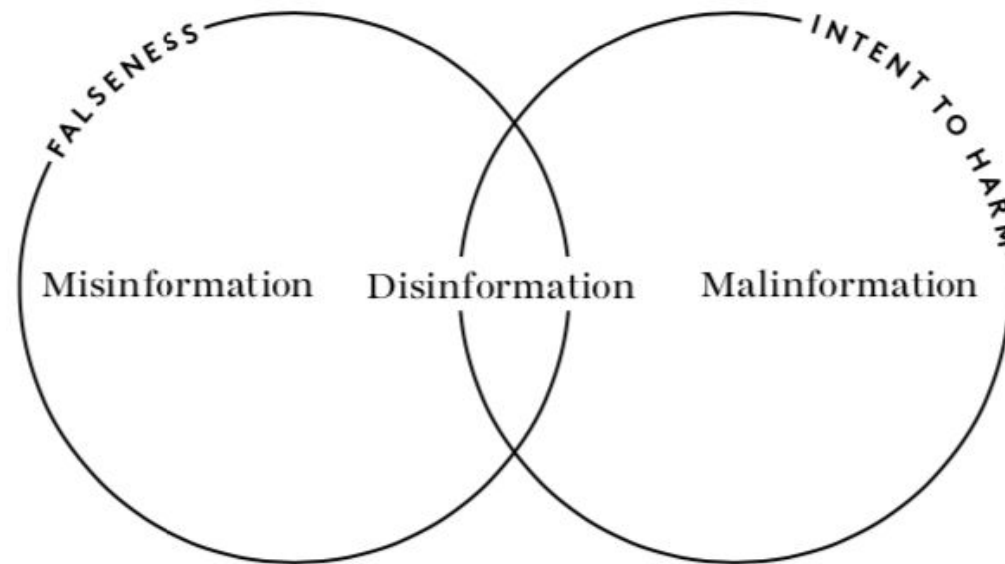
12/05/2025

Information Disorders

- In your group: What is the difference between misinformation and disinformation. Is this distinction important and if yes, why, if no, why not?

12/05/2025
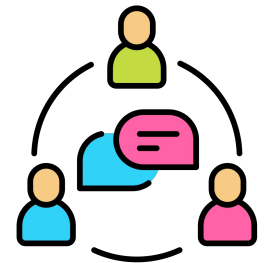
Information Disorders

- **Definition**
  - A concept describing illegitimate and harmful types of information, including **mis-**, **mal-**, and **disinformation**

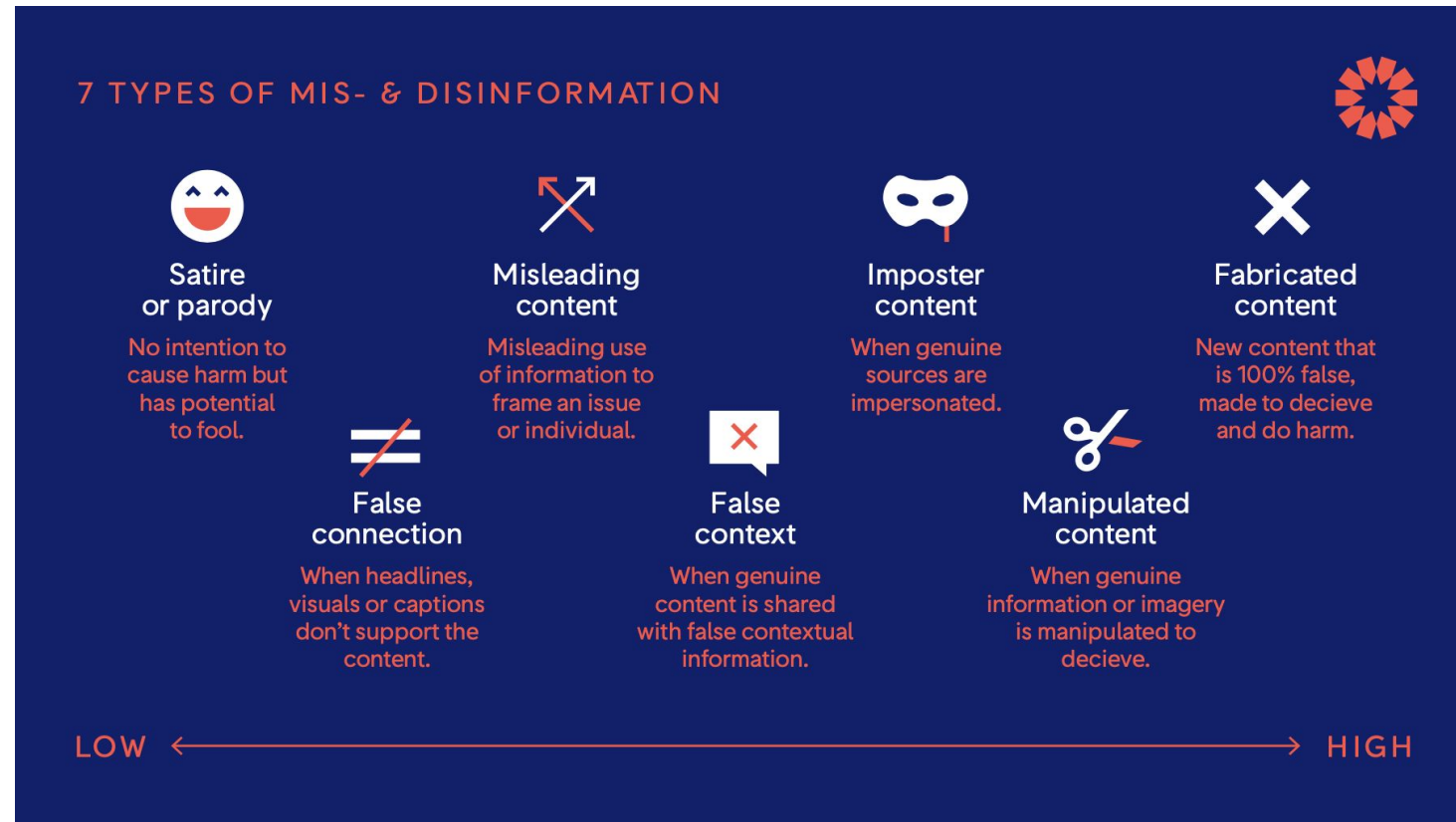Source: Wardle, 2020

Information Disorders

- **Definition**
  - Visualisation based on an outline of Keller et al., 2020. Dimensions: (1) Relationship towards truth and (2) intent

	True	Untrue
No intent to deceive	Information	*Mis*information
Intent to deceive	*Mal*information	*Dis*information

Source: Own illustration based on Keller et al., 2020

Types of mis- and disinformation

7 TYPES OF MIS- & DISINFORMATION

Satire or parody
No intention to cause harm but has potential to fool.

Misleading content
Misleading use of information to frame an issue or individual.

Imposter content
When genuine sources are impersonated.

Fabricated content
New content that is 100% false, made to decieve and do harm.

False connection
When headlines, visuals or captions don't support the content.

False context
When genuine content is shared with false contextual information.

Manipulated content
When genuine information or imagery is manipulated to decieve.

LOW ←——————————————————→ HIGH

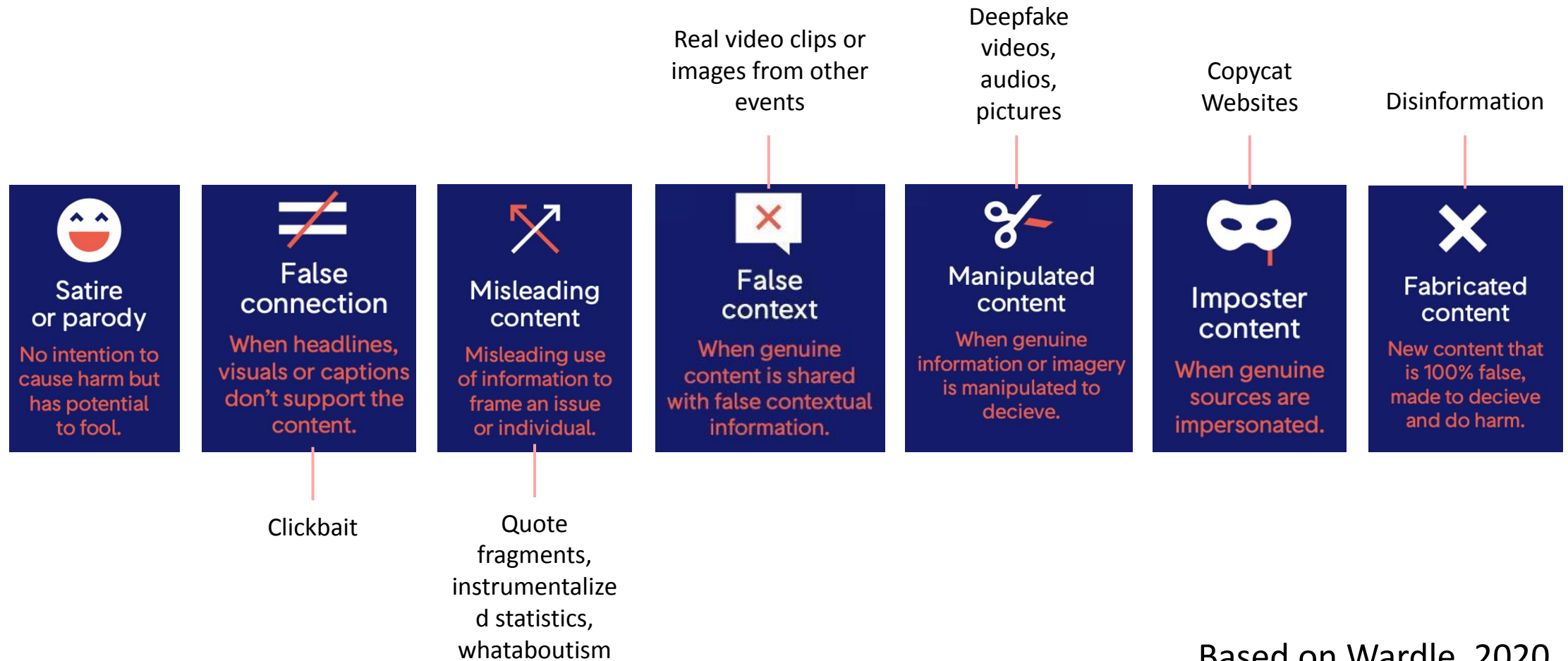Source: Wardle, 2020

12/05/2025

The spectrum of manipulation

- **Beyond disinformation…**
  - False news, false flag operations, inaccurate quotes, manipulated information, biased or misleading information, false amplification (Weedon et al., 2017)

Satire or parody	False connection	Misleading content	False context	Manipulated content	Imposter content	Fabricated content
No intention to cause harm but has potential to fool.	When headlines, visuals or captions don't support the content.	Misleading use of information to frame an issue or individual.	When genuine content is shared with false contextual information.	When genuine information or imagery is manipulated to decieve.	When genuine sources are impersonated.	New content that is 100% false, made to decieve and do harm.

Based on Wardle, 2020

12/05/2025

Informationsmanipulation umfasst Form und Inhalt

Real video clips or images from other events

Deepfake videos, audios, pictures

Copycat Websites

Disinformation

Satire or parody
No intention to cause harm but has potential to fool.

False connection
When headlines, visuals or captions don't support the content.

Misleading content
Misleading use of information to frame an issue or individual.

False context
When genuine content is shared with false contextual information.

Manipulated content
When genuine information or imagery is manipulated to decieve.

Imposter content
When genuine sources are impersonated.

Fabricated content
New content that is 100% false, made to decieve and do harm.

Clickbait

Quote fragments, instrumentalized statistics, whataboutism

Based on Wardle, 2020

12/05/2025

Disinformation

- **Goals**
  - Three motivations: make **money**; have **political influence**, cause **trouble** (for the sake of it) (Wardle, 2020)
- **On the diversity of disinformation motives**
  - In your group: Outline a plan for your own disinfo intervention to accomplish one of the goals above (to make money, gain political influence, cause trouble and chaos) and think of real-world examples of disinformation incidents for each objective

12/05/2025
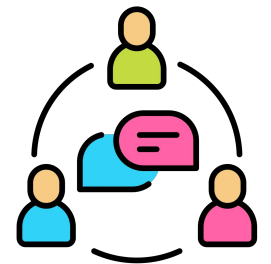
Disinformation

- **Example 1**: **Make money**
  - "More than 800,000 people in Europe and the US appear to have been duped into sharing card details [...] with a vast network of fake online designer shops apparently operated from China."

Source: Theguardian.com (link)

12/05/2025

Disinformation

- **Example 2**: **Political influence**
  - *Vote Leave* campaign bus in May 2016 used by the Johnson administration ahead of Brexit referendum

Source: Bloomberg.com (link)

12/05/2025

Disinformation

- **Example 3: Cause trouble**
  - Fake accounts: Twitter confirmed that the *Twitter* account @TEN_GOP was a fake, run by a Russian operative connected to a troll factory in St. Petersburg.

Tweet:

Tennessee @TEN_GOP   Aug 23

BREAKING: Antifa creating chaos in Phoenix. Gas is being deployed outside the convention center.

Source: medium.com - @DFRLab (link)

Tennessee
@TEN_GOP

12/05/2025

Disinformation

- **Note:**
  - Truly manufactured information and fabricated content is rather rare (Butler et al., 2024)
  - It is but one among many and more subtle attempts to manipulate public discourse (e. g., decontextualised content, partial content, doctored content, framing)

12/05/2025

Malinformation

- **Definition**
  - Strategic spread of true facts but with a negative intent (Keller et al., 2020)

- In your group: Which actions would fall under this definition? Do you know any real-world examples?
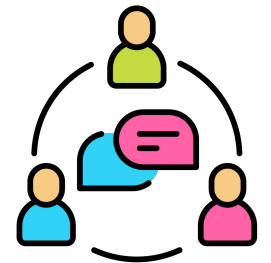
12/05/2025

Malinformation

- **Definition**
  - Strategic spread of true facts but with a negative intent (Keller et al., 2020)

- **Examples**
  - The leaking of emails of John Podesta, chairman of Hillary Clinton's 2016 presidential campaign
  - Cyberattack attributed to a Russian military intelligence agency. Whether the emails were doctored remains unclear. In case of doctoring, it would be disinformation

12/05/2025

Malinformation

- **Malinformation**
  - Strategic spread of true facts but with a negative intent (Keller et al., 2020)

- **Malinformation as a viewpoint dependent behaviour**
  - *WikiLeaks* also obtained the leaked material and published over 20.000 pages of emails (Carroll, 2016).
  - In your group: Would you consider the behaviour of *WikiLeaks* ethically legitimate or an act of malinformation? Note down possible arguments for both positions

12/05/2025
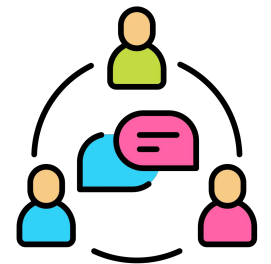
Misinformation

- **Definition**
  - False information spread without intention to harm (Keller et al., 2020)

targeting

Disinformation

Disinformation

unaware

aware

Misinformation

Disinformation

Source: Own illustration

12/05/2025
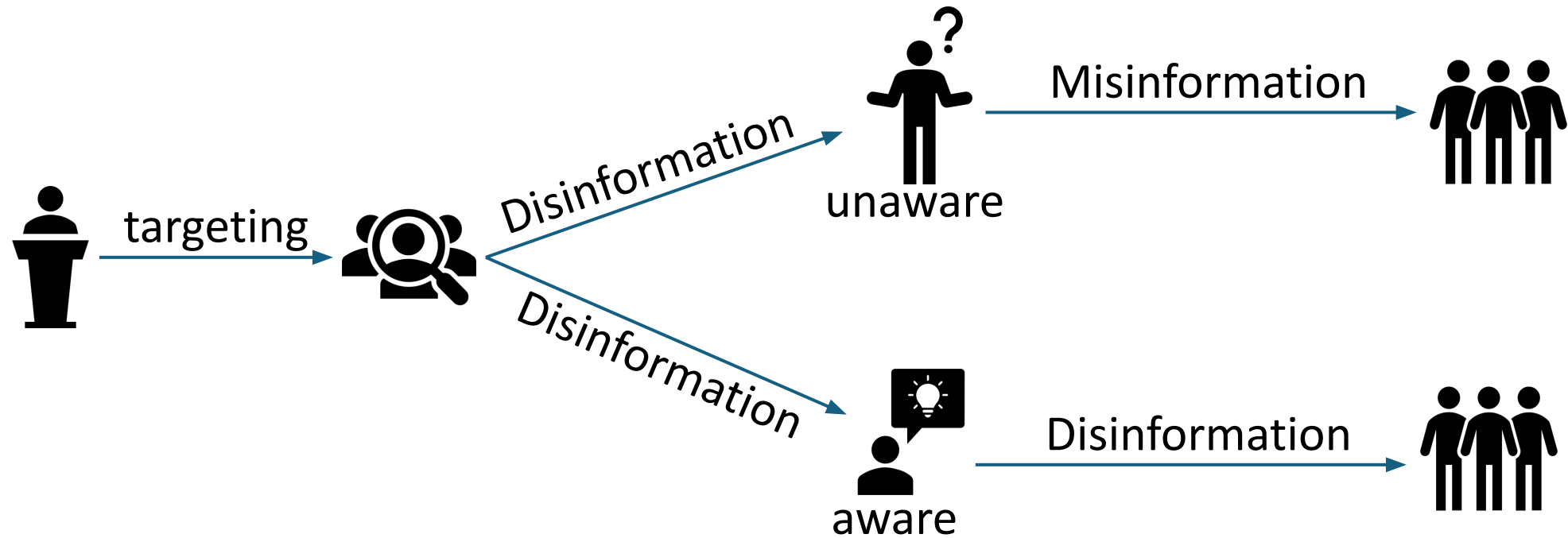
Fake News

- In your group: How would you define the term "Fake News" and how would an incident of Fake News look like?

- Although it is still being used in academic literature, there is an open discussion on whether to better drop the term. Can give an explanation why?

12/05/2025
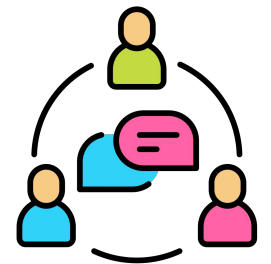
Fake News

- **Definition**
  - Still being used in academia, for example to describe forms of disinformation that is purposefully given a news-like appearance (George et al., 2021)
  - Term was co-opted and appropriated by Donald Trump (Coll, 2017) and other to denounce mostly moderate to left-leaning and public broadcasting media
  - In your group: Would you advocate to continue using the term "Fake News" in academia, or would you abandon it?

Source: newyorker.com
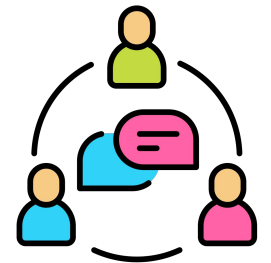
What is information manipulation?

- Definition
  - **Information Manipulation**: appealing to irrational impulses and using other communicative tactics to influence the beliefs and actions of target audiences (Klemp, 2010)
    - Share of truly **manufactured information** rather rare (Butler et al., 2024)

information manipulation

Source: own illustration

Literature

Armitage, R. & Vaccari, C. (2021). Misinformation and Disinformation. In H. Tumber & S. Waisbord (Eds.). The Routledge Companion to Media Disinformation and Populism. Routledge.

Butler, L. H., Prike, T., & Ecker, U. K. H. (2024). Nudge-based misinformation interventions are effective in information environments with low misinformation prevalence. Scientific Reports, 14(1). https://doi.org/10.1038/s41598-024-62286-7

Carroll, L. (2016). Are the Clinton WikiLeaks emails doctored, or are they authentic? Politifact https://www.politifact.com/article/2016/oct/23/are-clinton-wikileaks-emails-doctored-or-are-they-/

Coll, S. (2017). Donald Trump's "Fake News" Tactics. The New Yorker https://www.newyorker.com/magazine/2017/12/11/donald-trumps-fake-news-tactics

George, J., Gerhart, N., & Torres, R. (2021). Uncovering the Truth about Fake News: A Research Model Grounded in Multi-Disciplinary Literature. Journal of Management Information Systems, 38(4), 1067–1094. https://doi.org/10.1080/07421222.2021.1990608

Keller, F. B., Schoch, D., Stier, S., & Yang, J. (2019). Political astroturfing on Twitter: How to coordinate a disinformation campaign. Political Communication, 37(2), 256–280. https://doi.org/10.1080/10584609.2019.1661888

Klemp, N. (2010). When Rhetoric Turns Manipulative: Disentangling Persuasion and Manipulation. In W. Le Cheminant & J. M. Parrish (Eds.). Manipulating Democracy. Democratic Theory, Political Psychology, and Mass Media. Routledge.

12/05/2025

Literature

Wardle, C. (2020). Understanding Information Disorders. firstdraftnews.org.
https://firstdraftnews.org/long-form-article/understanding-information-disorder/

Weedon, J., Nuland, W., & Stamos, A. (2017). Information Operations and Facebook.
https://www.mm.dk/wp-content/uploads/2017/05/facebook-and-informationoperations-v1.pdf

World Economic Forum (2023, Aug. 4). Toolkit for Digital Safety Design Interventions and Innovations: Typology of Online Harms. (last retrieved: 2024, Sep. 12)
https://www.weforum.org/publications/toolkit-for-digital-safety-design-interventions-and-innovations-typology-of-online-harms/

Zerback, T., Töpfl, F., & Knöpfle, M. (2020). The disconcerting potential of online disinformation: Persuasive effects of astroturfing comments and three strategies for inoculation against them. New Media & Society, 23(5), 1080–1098.
https://doi.org/10.1177/1461444820908530

12/05/2025

Co-funded by the
European Union