

## MODULE:

- **Project:** 101132954—DRONE—ERASMUS-EDU-2023-PI-FORWARD
- Teacher and school leaders training to promote Digital literacy and combat the spread of disinformation among vulnerable groups of

## Technical Literacy to Combat Disinformation For Teachers

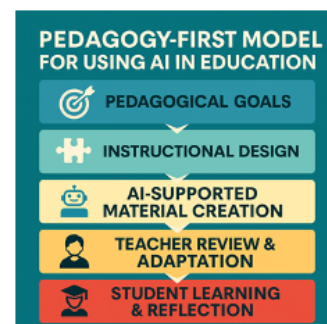
### ***READING for teachers***

**Created by the ELTE DRONE team (v.2025.12.12):**

- *Dr. Márta Turcsányi-Szabó, habil. associate professor  
([tszmarta@inf.elte.hu](mailto:tszmarta@inf.elte.hu)),  
the Dean's commissioner for educational innovation at ELTE  
Faculty of Informatics*

- *Franciska Mikófalvy, ELTE IK T@T Kuckó coordinator*

Using a pedagogy-first model, the material was designed and developed in collaboration with ChatGPT 5.0 Perplexity Pro as co-creator.



## CONTENT

<b>READ</b>	<b>3</b>
Introduction	3
Competencies for teachers	5
Adolescents need to acquire	8
<b>COURSE PATH</b>	<b>11</b>
I. Disinformation	13
II. The role of AI in dissemination and detection	17
III. Understanding how AI works and how to identify	19
1. Understanding how AI works & the ethics of AI	19
2. Uncovering algorithmic bias & fostering inclusivity	22
III. How does Cybersecurity awareness help?	24
1. Password puzzle: How your weak login becomes a weapon	24
2. Jeopardising security	26
3. Mastering platform safeguards to disrupt disinformation	28
4. Checking origin of email (phishing & disinformation)	31
6. Identifying bots	35
IV. Cultivating responsible Digital citizenship	37
1. Think before sharing	37
2. Netiquette use (Promoting respectful online communication)	40
3. Information hygiene	42
4. Echo Chamber	44
5. Understanding & managing your Digital footprint	46
6. The disinformation web	48
Adapting activities for adolescents: A teacher's guide	51
<b>KEY TOOLS+RESOURCES</b>	<b>54</b>

Disinformation is **misleading content that is deliberately created and spread to deceive people**, or to achieve economic or political gain, often with the intention of causing public harm. It is an orchestrated adversarial activity where actors use strategic deceptions and media manipulation tactics to advance political, military, or commercial goals.

### Key aspects of disinformation include:

- **Intentional and Strategic Nature:** Unlike misinformation, which is false information spread unintentionally, disinformation is purposefully disseminated with a clear intent to mislead or manipulate. Malinformation, another related term, refers to factual information shared with the intention to cause harm, often by taking it out of context or making private information public.
- **Tactics and Modalities:** Disinformation campaigns are implemented through coordinated efforts that "weaponize multiple rhetorical strategies and forms of knowing—including not only falsehoods but also truths, half-truths, and value judgements—to exploit and amplify culture wars and other identity-driven controversies". These campaigns can circulate through various deceptive behaviours and harmful content, including:
  - **Astroturfing:** Centrally coordinated campaigns that mimic grassroots activism.
  - **Conspiracy Theories:** Rebuttals of official accounts that propose alternative explanations involving secret groups or orchestrated events.
  - **Clickbait:** Deliberate use of misleading headlines and thumbnails for profit or popularity.
  - **Computational Propaganda:** Emphasizes the role of automation, such as bots and algorithms, to spread disinformation at scale.
  - **Internet Manipulation:** The use of online digital technologies, algorithms, social bots, and automated scripts for commercial, social, military, or political purposes.
- **Objectives:** The primary goals of disinformation are often to destabilize liberal democracies, undermine alliances, exacerbate societal divides to advance geopolitical agendas. It aims to create confusion and "information paralysis," leading people to distrust everything they encounter, making them easily manipulable.

- **Historical Context:** Disinformation is not a new phenomenon; it has been a feature of human communication since at least Roman times.
  - **Ancient Rome:** Octavian waged a propaganda campaign against Antony using coins with "short, sharp slogans" to smear his reputation.
  - **Printing Press Era:** The invention of the Gutenberg printing press dramatically amplified the dissemination of disinformation, leading to the first large-scale news hoax, 'The Great Moon Hoax' of 1835.
  - **20th Century:** The advent of radio and television further developed one-to-many communications, and propaganda played crucial roles in World War I and II, demonizing enemies and appealing to nationalism.
- **Modern Dangers (Digital Age):** The arrival of the internet and social media has **dramatically multiplied the risks and effectiveness of disinformation**.
  - **Connectivity and Computing:** High internet and social media usage, combined with advancements in AI technologies, allow actors to cheaply create and launch disinformation at an unprecedented scale and speed.
  - **Lack of Filters:** Unlike during the Cold War, there are currently few effective filters to stop disinformation from reaching users through the internet and social media.
  - **Erosion of Trust:** The sheer volume of information and repeated exposure to falsehoods can erode public trust in all sources, including legitimate news and democratic institutions.
  - **Challenges for Detection:** Disinformation is intentionally crafted to be difficult to detect, and its spread on social media generates large, incomplete, unstructured, and noisy data, making traditional detection algorithms ineffective. Anonymity also complicates tracing the origin and holding perpetrators accountable.

Overall, disinformation leverages human psychology and technological advancements to spread intentionally false narratives, posing significant threats to informed public discourse, democratic processes, and societal trust.

Technology and Artificial Intelligence (AI) play a multifaceted and increasingly critical role in both the **dissemination** and **detection** of disinformation.

## Competencies for teachers

The preparation of adolescents to combat disinformation is intrinsically linked to the competencies and principles outlined in the AI competency framework for teachers, as teachers are central to guiding students in the era of AI. The UNESCO AI competency framework for teachers (AI CFT) was, in fact, developed "hand in hand with a competency framework for students", signifying a coordinated approach to preparing both educators and learners.

Here's how the advice for preparing adolescents aligns with the AI competency framework for teachers:

- **Developing foundational Media literacy and Critical thinking skills:**
  - **Understanding AI's role:** The AI CFT aims to foster a **"Human-centred mindset"** among teachers, ensuring they understand that AI is "human-led" and that the creators' decisions significantly impact human rights and agency. Teachers should learn how AI is trained, and about data and algorithms. This knowledge enables teachers to **debunk AI hype** and explain to children that AI tools are products of human design, not infallible entities. Teachers need to convey that AI cannot replace human thinking or intellectual development.
  - **Asking critical questions:** Teachers, guided by the AI CFT, are encouraged to nurture **critical methodologies** for evaluating the benefits and risks of AI. The "Ethics of AI" aspect equips teachers with the skills and knowledge to understand and **critically evaluate AI tools, including their explainability and safety**. This directly translates into teachers being able to guide students in questioning information, understanding who created it, and why.
  - **Differentiating and verifying:** The AI CFT emphasizes that AI-generated content can be **stochastic and less trustworthy**, requiring teachers to critically examine its accuracy and design appropriate pedagogical methodologies. Teachers should be prepared to address the challenges posed by synthetic media like deepfakes and generative text, which make disinformation more convincing. Equipping teachers with skills to assess the **proportionality** and context-appropriateness of AI use is crucial for teaching students to discern valid information.
  - **Recognising emotional manipulation:** While not explicitly termed "emotional manipulation" in the AI CFT, the "Ethics of AI" aspect addresses **"algorithmic biases"** and the importance of **"promoting inclusivity"**. Disinformation often leverages emotional responses; teachers equipped to understand AI's societal impact and ethical concerns are better positioned to teach students how content can be designed to persuade or exploit their emotions. The AI CFT for students

also focuses on developing an understanding of AI's societal impact, including its potential to amplify biases.

- **Fostering human agency and ethical principles:** This is a core principle underpinning the entire AI CFT, which advocates for "protecting teachers' rights, **enhancing human agency**, and promoting sustainability". Teachers are trained to adopt a **"human-centred approach"** to AI, ensuring that technology serves to strengthen human capacities and promote human dignity. Teachers are seen as "guardians of safe and ethical practice" and "role models for lifelong learning about AI".
- **Cultivating responsible Digital citizenship:**
  - **Thinking before sharing and practicing "critical ignoring":** Teachers are expected to internalize essential ethical rules for the **safe and responsible use of AI**, including respecting data privacy, intellectual property rights, and avoiding the dissemination of disinformation or misinformation. This equips them to teach students the importance of verifying information before sharing and engaging critically with content, promoting a habit of "critical ignoring".
  - **Understanding group influence and engaging in social correction:** Teachers with a **"Human-centred mindset"** understand AI's societal impact and implications for citizenship. The AI CFT for teachers promotes their role in educating students on how to responsibly interact with online communities and address misinformation constructively, which aligns with teaching children about group influence and social correction.
- **Implementing educational strategies through schools and curricula:**
  - **Integrating Media literacy:** The AI CFT directly supports the integration of media literacy and critical thinking into school curricula. It provides a framework for designing **national AI competency frameworks and training programmes for teachers**, which can include specific modules on combating disinformation. It emphasizes that media literacy should be integrated "from an early age".
  - **Providing age-appropriate learning environments:** The AI CFT for students (developed alongside the teacher framework) highlights the importance of age-appropriate learning environments, including **"unplugged" activities** and **locally available AI tools**. Teachers equipped through the AI CFT are prepared to select and integrate appropriate AI tools into their pedagogical strategies.
  - **Utilising interactive tools:** The student framework suggests interactive methods like "[Play 'AI Bingo'](#)" or "Build a classifier in [Google's Teachable Machine](#)", which teachers, after gaining the "AI pedagogy"

competencies from the AI CFT, can effectively implement.

- **Teacher preparedness:** The entire AI CFT addresses the **urgent need for teachers to be empowered** to better understand the technical, ethical, and pedagogical dimensions of AI. It aims to support teachers in developing capabilities to leverage AI's benefits while mitigating its risks. Teachers need competencies to understand and critically evaluate AI tools and outputs, and design appropriate pedagogical methodologies to guide the use of AI-synthesized content.
- **Empowering parents in the process:**
  - While the AI CFT for teachers does not directly detail parent engagement, it lays the groundwork by preparing teachers to be **knowledgeable and ethical guides** in the AI era. A teacher who understands AI's implications and how to teach about it responsibly is better positioned to advise parents, foster open dialogue, and collaborate on supervising children's digital activity. This also ties into the need for a "whole-of-society" approach to combat disinformation, which includes parents.
- **Promoting a collective and collaborative approach:**
  - The AI CFT for teachers itself is a product of "contributions from a wide range of stakeholders". The framework highlights that a human-centred approach to AI requires regulators, AI providers, and institutions to be **co-responsible for governance**. This aligns with the broader "whole-of-society" approach needed to combat disinformation, fostering collaboration between governments, civil society, media, and technology companies.

In summary, the AI competency framework for teachers provides the essential blueprint for equipping educators with the values, knowledge, and skills necessary to prepare elementary school children to navigate the complex information landscape, particularly concerning AI-driven disinformation.



## Adolescents need to acquire

Adolescent children should learn a comprehensive set of skills and habits to prevent the spread of misinformation, focusing on media literacy, critical thinking, digital responsibility, and practical strategies for navigating today's complex information landscape. Here's what research and expert guidance recommend:

### 1. Media literacy and Critical thinking

- **Access and analyse information:** Learn to effectively access, analyse, and engage with media messages, understanding how to find, evaluate, and use credible information 1 2 3.
- **Ask key questions:** Always question the source, purpose, and credibility of information. Who created it? Why was it published? Is it verified? What's the format (article, meme, video)? 3
- **Compare multiple sources:** Check how different outlets or platforms report the same topic to spot inconsistencies or bias 3.
- **Recognize signs of fake news:** Be alert for overly sensational headlines, spelling mistakes, lack of credible sources, and content that's widely shared without verification 3.

### 2. Practical verification skills

- **Use fact-checking tools:** Learn to use resources like Google Images (for reverse image search), Decodex, or Hoaxbuster to verify images and stories 3.
- **Understand platform safeguards:** Be aware of built-in features on social media (like reporting, blocking, or flagging content) and use them to help control the spread of fake news 4.

### 3. Healthy scepticism and “critical ignoring”

- **Be sceptical, not cynical:** Maintain a healthy scepticism—don't automatically believe or share something just because it's popular or emotionally engaging. Learn to “critically ignore” suspicious content by not liking, commenting, or sharing it 4 5.
- **Limit information overload:** Recognize when to turn down the flow of online information and take breaks to avoid being overwhelmed by constant news and updates 5.

### 4. Responsible digital citizenship

- **Think before sharing:** Always verify information before sharing it with others. Understand the impact of spreading false information and strive to be a responsible participant in online communities 1 2 3.
- **Engage in social correction:** If you see friends or others sharing misinformation, learn how to address it respectfully and constructively 4.



## 5. Understand emotional and social influences

- **Be aware of emotional manipulation:** Recognize that fake news often plays on emotions like fear, anger, or excitement to go viral. Pause and reflect before reacting or sharing 3.
- **Group effect and peer pressure:** Understand that seeing information shared by many people doesn't make it true. Don't let peer pressure override your critical judgment 3.

## 6. Practice news evaluation and prebunking

- **Evaluate intentions:** Examine and evaluate the intention behind the media you consume and create. Is it meant to inform, shock, persuade, or sell? 2 3
- **Prebunking techniques:** Learn about common tactics used in misinformation (polarization, conspiracy theories, emotional appeals) through games or educational activities, which help build resistance to manipulation 6.

## 7. Open communication and self-reflection

- **Discuss online experiences:** Regularly talk with parents, teachers, or trusted adults about what you see online, and seek advice when in doubt 1 2 3.
- **Reflect on your own beliefs:** Question your assumptions and be open to changing your mind when presented with credible evidence 3.

## Summary table: What adolescents should learn

Skill/Strategy	Why It matters
Media literacy & critical thinking	Distinguish credible from false information
Verification tools	Confirm accuracy before sharing
Healthy scepticism	Avoid falling for viral or emotional manipulation
Responsible sharing	Prevent further spread of misinformation
Emotional/social awareness	Recognize manipulation and peer influence
Prebunking & evaluation	Build resistance to common misinformation tactics
Open dialogue & reflection	Seek guidance and be open to learning

## In summary:

Adolescents should develop strong media literacy, critical thinking, and digital citizenship skills. This includes learning to question sources, verify information, ignore or report suspicious content, and reflect on emotional and social influences. By practicing these habits, teens can help stop the spread of misinformation and become more resilient, responsible participants in the digital world<sup>1 2 4 5 3</sup>.

1. <https://parentandteen.com/how-to-teach-teens-to-navigate-misinformation/>
2. <https://marylandmedialiteracy.org/tweens-teens.html>
3. <https://www.softkids.net/en/esprit-critique-et-fake-news-comment-aider-les-enfants-a-sinformer-intelligemment/>
4. <https://www.nature.com/articles/s41599-024-04237-1>
5. <https://www.edweek.org/teens-are-digital-natives-but-more-susceptible-to-online-conspiracies-than-adults/2023/08>
6. <https://www.unicef.org/innocenti/media/856/file/UNICEF-Global-Insight-Digital-Mis-Disinformation-and-Children-2021.pdf>
7. <https://openarchive.tk.mta.hu/622/1/1-s2.0-S0747563224002061-main.pdf>
8. <https://www.internetmatters.org/issues/fake-news-and-misinformation-advice-how-protecting-children-from-fake-news/>
9. <https://www.apa.org/monitor/2024/09/media-literacy-misinformation>
10. <https://www.nytimes.com/2022/10/20/learning/lesson-plans/teenagers-and-misinformation-some-starting-points-for-teaching-media-literacy.html>

## INTRODUCTION (10 min)

### General objective of the module

To gain competencies in AI & cybersecurity for developing competencies in adolescents to become responsible digital citizens combating the spread of disinformation and learn practical strategies for navigating today's complex information landscape.

### Objectives for teachers:

To develop competencies to supervise and guide digital activities for adolescents to:

- Understand the concept of disinformation
- Understand how AI works and how to identify
- Understand how cybersecurity awareness can help
- Become responsible digital citizen
- Collaborate with schools and communities

### Objectives for students:

To gain competencies in AI & cybersecurity for adolescents to become responsible digital citizens combating the spread of disinformation and learn practical strategies for navigating today's complex information landscape.

To develop competencies for adolescents:

- Knowledge & Understanding: Concept of disinformation, AI/Cybersecurity awareness, Digital citizenship
- Use & Apply: Passwords, Emails & Web, Public Networks, Risks & Safety procedures
- Evaluate & Create: Risks & Safety, responsible citizenship
- Ethics: Risks & Safety of digital citizenship, AI awareness

### Personal introduction:

- Teacher Introduction
- Everyone should introduce themselves

### ADMINISTER ENTRY ASSESSMENT (15 min):

- PRE-TEST
- SELF-CHECK 
- Discuss *"Which one surprised you the most?"*

In case the course provides microcredentials, then a more comprehensive test could be administered before the course to allow participants to skip the module if successful (80% achieved):

## ACTIVITIES:

- *Disinformation (20 min)*
- *The role of AI in dissemination and detection (20 min)*
- *Understanding how AI works and how to identify (30 min)*  
*Active participation: 5 pts*
- *How does cybersecurity awareness help? (60 min)*  
*Active participation: 5 pts*
- *Cultivating responsible digital citizenship (60 min)*  
*Active participation: 5 pts*

### **Summary (20 min)**

*If necessary, commit some time to explore “specific detection tools.*

### **Exit assessment (10 min) 20 pts**

*Home essay: Read the file provided and compose a one pager describing the problem situation and background of your class and define an activity that would best suit the class as a solution for their age. Submit as Word.*



**Final Quiz: 100 pts**

**All together 140 points can be obtained.**

**80 points should be achieve in order to succeed.**

🕒 **OVERALL duration:** 4 hours in synchronous format + 1 hour for Exit assessment

# I. Disinformation

Go through the presentation of definitions and discuss issues.

**Definition:**

- **Disinformation:**  
*False information deliberately created and spread to deceive or mislead people.*  
*Example: Fake news articles designed to manipulate public opinion.*
- Key Characteristics of Disinformation**
- 1. Intentional deception**
  - The defining feature: **It is spread on purpose to mislead.**
  - Unlike misinformation (which can be unintentional), disinformation involves a deliberate strategy.
- 2. Manipulated content**
  - May use **doctored images, fake videos (deepfakes), or misleading headlines.**
  - Sometimes real facts are taken out of context to support a false narrative.
- 3. Emotional manipulation**
  - Often appeals to **fear, anger, or sympathy** to grab attention and go viral.
  - Designed to provoke strong reactions that override critical thinking.
- 4. Fabricated or hidden sources**
  - Disinformation may come from **fake accounts, bots, or websites** that mimic real news outlets.
  - It may hide who created or funded it.
- 5. Spread for gain**
  - Used to **influence public opinion, sow confusion, undermine trust, or achieve political, financial, or ideological goals.**
- 6. Often blends truth with lies**
  - May contain **some accurate information** mixed with falsehoods, making it harder to detect.
  - Known as “**malinformation**” when truths are presented in misleading ways to cause harm.

**Examples:**

- A fake tweet that appears to be from a public health organization, spreading false vaccine risks.
- A video edited to make a politician appear to say something they never said.
- **Phishing:**  
 Intentional deception: Phishing messages are crafted to trick recipients into sharing sensitive information—knowing they’re misleading.  
 False content: They usually impersonate legitimate entities (bank, school, platform) using fabricated URLs, fake branding, or bogus offers to mislead users.  
 Disinformation framework fit: Under standard definitions, phishing is an attempt to deceive for malicious gain—making it a form of disinformation, not malinformation  
**Common example:**

Type	Example
<i>Email phishing</i>	Fake email from your bank asking you to “verify your account.”
<i>Spear phishing</i>	Targeted message using your name, company, or role to seem more convincing.
<i>Smishing</i>	Phishing via SMS/text message.
<i>Vishing</i>	Voice call phishing, pretending to be tech support or a government agent.

**Emotional engineering:**

Examples include:

- Pretending to be tech support to get a password
- Tricking someone into clicking a malicious link
- Posing as a colleague to request sensitive data

- **Misinformation:**

*False or inaccurate information shared without intent to mislead.*

*Example: Sharing a wrongly captioned image believing it to be true.*

**Key characteristics of misinformation**

1. **Unintentional**

- The **person sharing it doesn't mean to mislead** others.
- It often comes from misunderstanding, poor fact-checking, or believing rumours.

2. **Based on false or incomplete information**

- May include **incorrect statistics, outdated facts, or misinterpreted research.**
- Sometimes it's simply **taken out of context.**

3. **Can appear credible**

- Misinformation often looks real or trustworthy, especially if it comes from friends, family, or familiar sources.
- **Spreads easily on social media** due to emotional or sensational content.

4. **May use real sources incorrectly**

- People may **misquote experts, use headlines without reading full articles, or rely on parody/satire without realizing it.**

5. **Can amplify harm**

- Even without intent, misinformation can **fuel panic, misguide decisions, or spread stereotypes.**
- **Repetition** can make false info feel "true" (the **illusory truth effect**).

**Examples:**

- Sharing a viral post about a natural remedy curing COVID-19, believing it's helpful.
- Reposting an outdated photo during a breaking news event, thinking it's recent.

- **Malinformation:**

*Genuine information used to harm a person, organization, or country.*

*Example: Leaking private data to embarrass someone.*

**Key characteristics of malinformation**

1. **Based on Genuine Information**

- The **content is factually correct**, but how and why it's shared is harmful.
- Example: Leaking someone's private messages or personal data.

2. **Intent to harm**

- The goal is to **damage reputations, incite hate, manipulate, or cause conflict.**
- Often used in political or personal attacks.

3. **Out of context**

- Real information is **presented without necessary background, used misleadingly, or distorted.**
- Example: Quoting someone accurately but removing context to reverse the meaning.

4. **Invasion of privacy**

- Sharing **personal or sensitive content**—like medical records, addresses, or images—without permission.
- Sometimes includes **doxxing** (publishing private information online).

5. **Timed for maximum damage**

- Malinformation may be **released strategically**, like during elections, protests, or public crises, to amplify fear or confusion.

**Examples:**

- Publishing hacked emails from a politician right before an election.
- Releasing a truthful news report about a person's past to ruin their reputation years later.
- Posting real photos of someone with a misleading caption to incite hate.

- **Fake news:**

*Fabricated news stories presented as legitimate journalism, often shared on social media.*

*Example: Viral headlines claiming false scientific discoveries.*

**Key characteristics of fake news**

1. **Entirely or Partially False**
  - The core content is **fabricated**, **distorted**, or **taken out of context**.
  - May use **clickbait headlines**, made-up quotes, or invented events.
2. **Disguised as real news**
  - Presented with the **appearance of professional journalism** (e.g., fake news sites mimicking real media outlets).
  - Uses **trusted formats** (logos, bylines, or datelines) to trick readers.
3. **Intentionally deceptive**
  - Created to **mislead**, often for **political gain**, **financial profit** (through ad revenue), or **social influence**.
  - Falls under **disinformation**, but specifically mimics journalism.
4. **Emotionally charged**
  - Uses sensational, shocking, or outrageous content to provoke strong feelings (fear, anger, outrage).
  - Drives engagement (clicks, shares, comments) without regard for truth.
5. **Lacks credible sources**
  - Often **no cited evidence**, **anonymous sources**, or **links to unreliable websites**.
  - Avoids fact-checkable claims or misrepresents studies/data.
6. **Designed to go viral**
  - Structured to be **easily shareable**, especially on social media.
  - Headlines may be **misleading or exaggerated** to maximize spread.

**Examples:**

- A fabricated article claiming a celebrity endorsed a miracle cure
- A fake news site publishing a false report that a politician was arrested
- A viral image with a completely made-up caption that contradicts the visual evidence

- **Deepfake:**

*Synthetic media (usually video or audio) generated by AI to imitate real people, often convincingly.*

*Example: A video showing a politician saying things they never actually said.*

**Key characteristics of deepfakes**

They are powerful forms of **synthetic disinformation** and can be highly realistic, making them difficult to detect without careful analysis or specialized tools.

1. **AI-generated or AI-manipulated**
  - Created using **machine learning techniques**, especially **deep learning** and **generative adversarial networks (GANs)**.
  - Often involves mapping one person's face or voice onto another's.
2. **Hyper-realistic appearance**
  - Designed to **look and sound convincingly real**, often mimicking facial expressions, lip movements, and voice tone with high precision.
  - Hard to spot with the naked eye.
3. **Deceptive purpose**
  - Frequently used for **malicious intent**, including:
    - o Political manipulation
    - o Celebrity hoaxes
    - o Fraud and scams
    - o Revenge or harassment
  - They can spread **false narratives** that damage reputations or mislead the public.



#### 4. Audio and visual manipulation

- Not limited to videos—can include **fake audio clips** of people saying things they never said.
- Voice cloning software can imitate tone, rhythm, and accent.

#### 5. Often shared on social media

- Deepfakes are typically **shared virally** through social platforms, gaining traction due to their shock or entertainment value.
- Many users may not realize they're fake.

#### 6. Fuel the “Liar’s dividend”

- Even real videos can be **falsely dismissed as deepfakes**, leading to **public distrust** in authentic evidence.
- This erodes confidence in video/audio as reliable forms of proof.

#### Examples:

- A fake video of a world leader making a provocative statement they never made
- An altered interview where a celebrity appears to confess to something untrue
- An AI-generated voice impersonating a CEO to commit financial fraud

### • Propaganda:

*Propaganda is information—often biased, misleading, or emotionally charged—used to influence people’s opinions, beliefs, or actions in favour of a particular cause, political agenda, or ideology.*

*Example: A government poster showing only the positive outcomes of a war effort while ignoring civilian casualties is a form of propaganda*

#### Key characteristics of propaganda:

- **Selective use of facts:** Uses facts, images, or quotes out of context to shape a message.
- **Emotional appeal:** Aims to trigger strong emotions like fear, pride, anger, or guilt.
- **One-sided messages:** Focuses only on one perspective and often ignores counterarguments.
- **Repetition:** Repeats messages or slogans to make them more memorable and persuasive.
- **Targeted persuasion:** Often directed at specific groups to influence their behaviour or beliefs.

#### Teacher activity:

- Provide examples for each.
- Group discussion: “Have you ever been misled by a piece of news online?”

#### Modern types:

- Deepfakes (AI-generated images/videos)
- Troll farms and coordinated campaigns
- Fake social media accounts and bots
- Synthetic text (AI-generated articles/comments)
- COVID-19 vaccine disinformation
- Financial scams using AI voices
- Election interference via social media

## II. The role of AI in dissemination and detection

Technology and AI don't just help stop fake information—they also make it easier for disinformation (false or misleading information) to spread quickly and widely. Here's how:

**Speed and Scale of Dissemination:** Fake news can now spread fast and far online—especially on social media—much faster than in traditional newspapers or TV. It can go viral in minutes with no one checking if it's true.

**Algorithmic Amplification:** Social media sites use algorithms to decide what you see. These algorithms often boost emotional or shocking content, even if it's false. This can create “echo chambers”—spaces where people mostly see opinions they already agree with, making it harder to spot disinformation.

**Hyper-Realism and Synthetic Media (Deepfakes):** AI can now create real-looking fake videos, images, or audio—these are called deepfakes. For example, someone could make a fake video of a person saying something they never said. This makes it hard to tell what's real and what's not. Some people even use the idea of deepfakes to claim real videos are fake—this is called the “liar's dividend.”

**Personalisation and Hyper-Targeting:** With all the data people share online, AI can target people with specific messages that match their beliefs or fears. This hyper-targeting makes disinformation feel more convincing and harder to ignore.

**Anonymity and Low Cost:** It's cheap and easy to spread lies anonymously online. Bad actors don't need much money or effort to start a disinformation campaign, and they can hide their identity easily.

**Troll Farms and Computational Propaganda:** Some groups—called troll farms—use fake accounts (sock puppets) and automated bots to post the same messages again and again. This can manipulate public opinion and pretend there's a big crowd supporting an idea. When these fake campaigns pretend to be real grassroots movements, it's called astroturfing.

**Exploiting Online Advertising Technologies:** Fake news websites can make money through ads. So there's a financial reason for people to create and share disinformation—the more people click, the more money they make.

Given the escalating threat, digital technologies and AI are also being developed and deployed to detect and prevent the spread of disinformation.

Technology and artificial intelligence (AI) are being used to help spot and stop fake news and false information online:

**Smart Programs Read Content:** Computers can now read and understand language. They look for clues in stories that might mean the information is false.

**Learning from Examples:** These tools get better by looking at lots of examples of real and fake news. Over time, they learn how to tell the difference.

**Instant Warnings:** Some tools work in real time, giving alerts or showing warning labels when something seems suspicious.

**Checking Where Content Came From:** Tools can now check the original source of a photo, video, or article. This helps make sure it hasn't been changed or faked.

**Looking at Details:** AI can examine tiny errors in images or videos, or even check behind-the-scenes information (like when or where it was made) to see if something's off.

**Helping Platforms Respond:** Social media sites use these tools to hide false content, remove ads from fake stories, or show fact-check labels.

**Tools for Users:** There are simple tools for teachers, students, and parents to help check if something is real—right in your web browser!

**Spotting Bots:** AI can find fake accounts that try to spread lies by copying and repeating messages.

In short, while technology can be used to spread fake stuff, it's also being used to fight back and keep people informed—especially when combined with good teaching and media skills.

### Specific AI detection tools to try out:

- **Detecting-AI.com:** A free online tool that uses multiple algorithms to detect whether a piece of writing was generated by AI (e.g., ChatGPT or other models). Version 2 offers improved accuracy, but is still not perfect!
- **Originality.AI:** A professional-grade AI detector used by educators and content creators. It checks for both **AI-generated content** and **plagiarism**. Paid plans available, often used in academic or publishing.
- **Hive AI – Deepfake Detection:** A powerful tool used by media and content platforms to identify **deepfake videos and manipulated media**. Hive AI uses computer vision to flag suspicious content in real time.
- **Sensity AI:** Specializes in detecting **synthetic media** (deepfakes) across video, image, and audio formats. Used by governments, journalists, and companies to monitor and prevent media manipulation.

### III. Understanding how AI works and how to identify

**Proposed activity:** Depending on the pre-test results of students choose the most relevant activity for the class.

#### 1. Understanding how AI works & the ethics of AI

**Duration:** 30 minutes (Choose from activities depending on background of teachers)

**Format:** Collaborative workshop

**Group size:** Small groups of 3–5

#### Objectives

- Experience how AI “learns” and makes decisions.
- Understand AI is not autonomous—it reflects human inputs, training data, and limitations.
- Explore ethical implications: bias, misinformation, responsibility.
- Build strategies for helping students critically assess AI-generated content.

#### Materials & tools

- Internet-enabled devices (1 per group)
- Access to:
  - [This X Does Not Exist](#) (image generator)
  - Reverse image search (Google), **Bad news game:** <https://www.getbadnews.com>
  - [ChatGPT](#) or [Bing Copilot](#) (text generator)
  - [Deepfake Video Examples](#) (or curated YouTube clips)
- Handouts:
  - [II.1.How AI Learns Simulation Activity.docx](#)
  - [II.1.AI Ethics Reflection Prompts.docx](#)
  - [II.1.AI Ethics Detective Worksheet.docx](#)

#### Workshop flow (Optional activities)

Phase	Time	Activity
1. Hook & starter discussion	5 min	Show AI-generated face (from <i>This Person Does Not Exist</i> ) and ask: “Real or fake?” Discuss how convincing it is and why.

<b>2. AI learning simulation (Analogy game)</b>	10 min	Use “How AI learns” Simulations Cards”: Each group trains a “classifier” (human) with example cards (e.g., fruit/not fruit). Then, give ambiguous inputs. Discuss: How did training influence decisions?
<b>3. Tool exploration rotation</b>	10 min	In groups, teachers try: (a) text generation (ChatGPT), (b) image generation, (c) deepfake example. Use a worksheet to assess credibility, signs of fakeness, and potential misuse.
<b>4. Ethics gallery walk</b>	10 min	Display ethical dilemma prompts (e.g., “AI-generated essays,” “Deepfake politician,” “Biased hiring tools”). Groups walk, read, and discuss each. Leave sticky-note questions/responses.
<b>5. Whole-group discussion &amp; debrief</b>	10 min	Discuss: “What surprised you?” “What should students know about these tools?” “Where is the human responsibility?”
<b>6. Exit reflection</b>	5 min	Each teacher writes one action to try in their teaching, and one concern they’ll continue thinking about. Option to share aloud.

## Discussion prompts

- “Can AI be creative—or just remix what it’s seen?”
- “Who decides what AI learns from?”
- “How do you know if a student used AI—and does it matter?”
- “If an AI makes a harmful decision, who is accountable?”

## Assessment rubric (Self & peer reflection)

Criterion	1 – Emerging	3 – Developing	5 – Proficient
Understanding of AI	Vague or confused	Basic grasp of pattern recognition and training	Clearly explains model logic, limitations
Ethical Awareness	Limited or naïve view	Recognizes some ethical issues	Insightfully explains bias, misuse, accountability
Collaboration	Passive or disengaged	Participates with some input	Shares ideas, listens, leads inquiry
Application to Teaching	No classroom ideas shared	Basic integration idea	Specific strategy for guiding students

## Extension options

- How AI works: <https://www.youtube.com/watch?v=xj3yYLinf0>

- Neural Networks - Machine Learning for Kids:  
<https://machinelearningforkids.co.uk/> How to use:  
<https://www.youtube.com/watch?v=-lCezpMs3tk>
- Tinker with Neural Networks:  
<https://playground.tensorflow.org/#activation=tanh&batchSize=10&dataset=circle&regDataset=reg-plane&learningRate=0.03&regularizationRate=0&noise=0&networkShape=4,2&seed=0.11829&showTestData=false&discretize=false&percTrainData=50&x=true&y=true&xTimesY=false&xSquared=false&ySquared=false&cosX=false&sinX=false&cosY=false&sinY=false&collectStats=false&problem=classification&initZero=false&hideText=false>
- Analyse real student work vs. AI output: How can you tell?
- Create classroom norms for responsible AI use.
- Debrief generative AI use in digital citizenship education.
- Google's Teachable Machine <https://teachablemachine.withgoogle.com>  
train a simple image or sound classifier (e.g., "smile vs non-smile").

Reflection Worksheet: [II.1.Ext.AI Myths vs. Reality Worksheet.docx](#)

Discuss:

- "What did you teach the model?"
  - "What mistakes did it make?"
  - "Why can't AI think like humans?"
- Reverse image search (e.g., Google Images)

## 2. Uncovering algorithmic bias & fostering inclusivity

### Objective

Analyse how biased datasets and non-inclusive AI systems can generate false narratives, skew perceptions, and inadvertently spread disinformation. They'll learn to identify bias, test inclusive models, and design inclusive pedagogy.

### Materials & tools

- **Computers/tablets** (1 per pair)
- **Gender shades demo by Joy Buolamwini** – (tests face recognition bias)
- **ChatGPT** (free) – <https://chat.openai.com>
- **Worksheet: “Bias Detective”** (see below)
- **Padlet or Miro** for insights
- Optional article for reflection: “[Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification](#)” by Buolamwini (PDF available online)

### 30-Minute Structure

1. **Brief Introduction (4 min)**
  - Facilitate a quick discussion: *What if AI mistakes a person's identity or misconstrues community messages? How could this fuel false beliefs or disinformation?*
2. **Hands-on bias testing (12 min)**
  - In pairs, upload a few diverse selfies (or sample photos) onto the **Gender Shades** tool.
  - Record the model's gender detection accuracy and time.
  - Prompt ChatGPT for an answer to a culturally nuanced question (e.g., “Name a cultural festival in India”). Check for errors, omissions, or overgeneralizations.
3. **Worksheet reflection (8 min)**
  - Use the “**Bias detective**” **worksheet** to capture:
    - Model tested
    - Example input
    - Error type (misclassification/omission)
    - Potential impact for disinformation
    - Inclusive fix or pedagogical action
4. **Group sharing & discussion (5 min)**
  - Post key findings (errors + disinformation risk) on Padlet/Miro.
  - Discuss how biased predictions might reinforce stereotypes or create misleading content.
5. **Wrap-up & action (1 min)**
  - Emphasize: *AI reflects the data it's trained on—so it can misinform in biased ways.*
  - Highlight inclusive AI strategies and curriculum integration.



## Worksheet: Bias detective

AI tool	Input (photo / query)	Outcome / Error	Potential for disinformation (e.g., mislabel community)	Inclusive Fix / teaching strategy
Gender shades	e.g. Person A (dark-skinned)	Misclassified gender; took longer to detect	Could reinforce stereotypes, misidentify identities	Teach dataset bias; use inclusive image banks
ChatGPT	e.g. "Name a girl's science hero from Africa."	Omitted lesser-known figures	Suggests African women aren't scientists—erases representation	Encourage diverse dataset prompts; student-led input

(Complete at least 2 entries)

## Assessment rubric

Criteria	3 pts (Excellent)	2 pts (Good)	1 pt (Needs Improvement)
<b>Bias identification</b>	Pinpoints clear bias/misclassification	Notes bias but lacks detail	Misses or misidentifies bias
<b>Disinformation risk</b>	Explains how bias creates false narratives	Mentions risk but without clear link	Rarely connects to disinfo danger
<b>Inclusive solution</b>	Proposes feasible dataset/teaching strategy	Suggests solution but lacks depth	No practical corrective approach
<b>Collaboration</b>	Actively shares, provides clear insights	Some participation, average clarity	Minimal or no peer engagement

- Reverse image search (e.g., Google Images)
- [Sensity AI's "Deepfake Detection Challenge"](#) (demo version)

## References:

<https://www.cyberdefensemagazine.com/the-illusion-of-truth-the-risks-and-responses-to-deepfake-technology/>

<https://researchmgt.monash.edu/ws/portalfiles/portal/668653278/640947365-oa.pdf>

[https://www.researchgate.net/publication/378180889\\_Assessment\\_framework\\_for\\_deepfake\\_detection\\_in\\_real-world\\_situations](https://www.researchgate.net/publication/378180889_Assessment_framework_for_deepfake_detection_in_real-world_situations)

### III. How does Cybersecurity awareness help?

**Proposed activity:** Depending on the pre-test results of students choose the most relevant activity for the class.

#### 1. Password puzzle: How your weak login becomes a weapon

##### Objectives:

- Understand the principles of strong password creation and management.
- Recognize how poor credentials can lead to account compromise and be exploited in **disinformation campaigns**.
- Explore strategies and classroom activities to teach students **password hygiene** and **account responsibility**.

**Duration: 60 minutes**

##### Materials:

- Laptops/tablets (1 per pair)
- [III.1.Password Health Checklist.docx](#)
- [III.1.Password Scenario Cards.docx](#)
- Website for password strength testing (e.g., <https://howsecureismypassword.net> or <https://nordpass.com/secure-password/>)
- [Have I Been Pwned](#)
- [Cybersecurity for Teachers \(Common Sense\)](#)
- [Google's Password Manager Tool](#)
- Sticky notes, markers, chart paper/Miro board for online sync

##### Lesson Flow:

Phase	Time	Activity
<b>1. Hook: The hacked teacher</b>	10 min	Present a fake email claiming the teacher's school account has been hacked. Ask: <i>"How did this happen?"</i>
<b>2. Background input</b>	10 min	Briefly explain: - Common password vulnerabilities - Credential stuffing & reuse - Link to disinformation (e.g. social accounts hijacked to spread fake news)
<b>3. Scenario puzzle game</b>	15 min	In pairs, teachers receive <b>"How It Was Hacked"</b> scenario cards and match them to poor practices (e.g.,

password reuse, name+birthyear). Share out one fix per group.

<b>4. Password health check</b>	10 min	Teachers review 1–2 of their own passwords (not shared aloud) using a strength checker. Use the <b>Password Health Checklist</b> to self-assess: length, complexity, reuse.
<b>5. Teaching brainstorm</b>	10 min	Small groups brainstorm how to teach password hygiene at various grade levels (e.g., password art, password vaults, gamified challenges). Chart and share ideas.
<b>6. Reflection &amp; exit note</b>	10 min	Sticky-note exit prompt: <i>“One thing I will stop, change, or start doing about my passwords.”</i> Share reflections.

### Discussion prompts:

- “What makes a password easy to guess?”
- “Why might students reuse passwords—or share them?”
- “How does poor password management make someone vulnerable to disinformation campaigns?”
- “What’s one myth about passwords we should stop teaching?”

### Assessment rubric (Formative – for facilitation reflection)

Criterion	1 – Beginning	3 – Developing	5 – Proficient
<b>Engagement in scenario matching</b>	Passive	Participates with guidance	Explains and justifies password weaknesses clearly
<b>Use of password health checklist</b>	Minimal or incomplete	Completes with some self-reflection	Completes thoughtfully and applies changes
<b>Collaboration &amp; brainstorming</b>	Few contributions	Shares 1 idea or strategy	Actively contributes multiple relevant classroom ideas
<b>Exit reflection</b>	Generic or vague	Reasonable, practical action	Specific, insightful takeaway with intended impact

### Optional extension:

Invite IT staff to show how accounts are protected in your school, or run a **“Create-a-Classroom-Password-Policy”** mini-project in follow-up PD.

## 2. Jeopardising security

**Theme:** *Protecting personal data online*

**Duration:** 60 minutes

**Group size:** Pairs or small groups (3–4)

### Learning objectives

By the end of the session, teachers will:

- Identify how online tools (quizzes, forms, games) gather personal data—intentionally or not.
- Recognize how such data can be exploited for disinformation or phishing attacks.
- Develop practices and classroom strategies for promoting safe digital behaviour.

### Materials & tools

- Laptops/tablets (1 per pair or group)
- Access to: [Quizizz](#) or [Kahoot!](#) or Google Forms or [Typeform](#) (*optional*)
- [III.2.What You Share vs Who Sees It.docx](#)
- [III.2.Reflection and Action Sheet.docx](#)
- Projector and whiteboard for discussion prompts

### Lesson flow

Phase	Time	Activity
1. Hook & awareness	5 min	Icebreaker quiz on Kahoot (fake or real quiz questions like “Your Spirit Animal” or “Which Hero Are You?”). After the quiz, pose: <i>What personal data did you just share—even hypothetically?</i>
2. Analysis: “Build a trap”	10 min	In pairs, teachers design their own <b>3-question quiz</b> (e.g., “What’s your pet’s name?”, “Birth month?”, “First school?”). Record it on the handout and analyze what personal identifiers could be inferred.
3. Swap & simulate	10 min	Groups swap quizzes and take them. Then analyze what <i>they learned about each other</i> from just the answers.
4. Disinformation link	10 min	Discussion: <i>How could this information be used to guess passwords? To customize a phishing email?</i> Introduce the idea of microtargeted disinformation campaigns based on quiz-leaked data.

5. Group reflection	10 min	Complete <b>“What You Share vs. Who Sees It”</b> chart together. Classify what data was collected and who might have access (e.g., platform, advertisers, hackers).
6. Whole-group debrief	10 min	What surprised you? What practices should we model or teach? Generate a list of <b>“3 Golden Rules”</b> for sharing data online.
7. Personal action plan	5 min	On the <b>Reflection Sheet</b> , each teacher writes 1 change they will make and 1 thing they’ll teach students next week.

### Discussion prompts

- *Would your quiz reveal answers to common password-reset questions?*
- *Why do these platforms collect this information?*
- *What does “free” really cost in online apps or games?*
- *How can this connect to identity theft or disinformation?*

### Assessment rubric: Awareness & application

Criteria	1 – Beginning	3 – Developing	5 – Proficient
<b>Data awareness</b>	Vaguely identifies shared data	Points out common personal data types	Analyses intent and risk of each data point shared
<b>Quiz analysis</b>	Creates basic quiz	Includes questions with some personal relevance	Designs questions that mimic real-world phishing techniques
<b>Security reflection</b>	Lists general tips	Applies learning to teacher role	Articulates changes to teaching or classroom digital hygiene policies
<b>Collaboration</b>	Minimal group interaction	Participates in planning and discussion	Actively contributes, reflects, and challenges group assumptions constructively

### Summary

This session helps teachers **experience the real risks** of data exposure through “fun” online interactions and **connects that awareness to the broader context of digital security, disinformation, and ethical AI use**. It fosters **collaboration, critical reflection, and practical classroom takeaway strategies**—core to the AILit Framework’s vision of AI literacy in education.

### 3. Mastering platform safeguards to disrupt disinformation

**Duration:** 60 minutes

**Group Size:** 3–5 per group

**Setting:** In-person or hybrid workshop with access to computers/tablets

#### Learning objectives

- Understand how to **identify and respond** to misinformation/disinformation on social media and messaging platforms.
- Learn to use **platform safeguards** (e.g., reporting, blocking, flagging).
- Practice integrating **these tools into digital literacy instruction** with students.

#### Materials & tools

- Devices with internet access (1 per group)
- [III.3.Platform\\_Safeguard\\_Cheatsheet.docx](#)
- [III.3.Disinformation\\_Scenario\\_Cards.docx](#)
- Projector/screen for platform walkthroughs
- Optional: Demo accounts or screenshots of reporting/blocking flows

#### Lesson flow

Phase	Time	Activity description
1. Hook & poll	5 min	Begin with a live poll (e.g., Mentimeter): “Have you ever reported something online? What stopped you if not?” Discuss barriers (confusion, doubt, fear).
2. Mini-demo	10 min	Facilitator demonstrates reporting/blocking features across 2–3 platforms using walkthroughs/screenshots.
3. Scenario challenge	15 min	Groups receive <b>Scenario Cards</b> : e.g., student sees conspiracy video; classmate shares fake giveaway; WhatsApp chain message. Identify platform, tools.
4. Plan & practice	15 min	Groups use the <b>Disinfo Response Planning Sheet</b> to decide: What tool to use? What would we teach a student to do here? What are the risks?

<b>5. Gallery walk</b>	10 min	Groups display their sheets. Others leave sticky-note feedback: "Would this work for your class?" "What might be missing?"
<b>6. Reflection &amp; action</b>	5 min	Close with discussion: "What surprised you?" "What's one way you'll build this into your teaching?"

## Discussion prompts

- "Why might students hesitate to report harmful content?"
- "What are the risks of ignoring a false claim?"
- "How can we help students think critically about platform feedback loops?"

## Worksheet: Safeguard strategy planner

Scenario description	Platform feature used	Steps taken	Expected outcome	Classroom teaching tip
e.g., Phishing DM	Report & block	Click "Report message"; block sender	Reduces risk and notifies platform	Teach students to screenshot and report with adult help
e.g., Hateful comment	Report + filter words	Report comment, mute key offensive words	Content hidden; moderator alerted	Use "hide word" feature to model self-care online

## Assessment Rubric

Criterion	1 – Emerging	3 – Developing	5 – Proficient
<b>Tool Identification</b>	Unclear which platform or tool to use	Recognizes correct platform, some uncertainty	Accurately matches platform features to disinfo situations
<b>Pedagogical Application</b>	Vague or missing student strategies	Proposes a basic teaching approach	Outlines clear, age-appropriate classroom integration
<b>Scenario Analysis</b>	Oversimplifies or misses risks	Identifies one or two key risks	Thoughtfully considers impact, ethical issues, and outcomes
<b>Collaboration</b>	Minimal group participation	Participates, some shared input	Fully engaged, co-creates solution, constructive feedback

## Optional add-ons



- Create a **student-facing version** of the response template.
- Host a “Simulated Platform Lab” with mock disinformation content.
- Invite a guest (e.g., platform safety expert or journalist).

## Key takeaways

- Teachers gain **hands-on experience** using platform safeguards like reporting, blocking, flagging, and filtering.
- They understand **why and how** these tools curb disinformation and harm.
- Built-in features reflect educators’ roles in modelling **responsible digital citizenship** with agency, aligned with UNESCO’s AI CFT principles.
- Learning to apply these in classrooms equips students with essential skills for ethical and informed online engagement.

## 4. Checking origin of email (phishing & disinformation)

### “Inbox detective: Tracing the truth”

**Duration:** 60 minutes

#### Objective:

- Understand the structure of phishing emails
- Identify key signs of disinformation and manipulation in email form
- Practice tracing email sources and URLs
- Learn strategies to teach students safe and skeptical email habits

#### Materials needed:

- [III.4.Email\\_Red\\_Flags\\_Checklist.docx](#)
- [III.4.Inbox\\_Detective\\_Email\\_Pack.docx](#)
- Devices with internet access (1 per pair)
- Optional tool: [Google's Phishing Quiz](#)
- Whiteboard or digital collaboration board (Jamboard, Padlet)

#### Lesson flow

Phase	Time	Activity
<b>1. Hook – email from a “prince”</b>	10 min	Project a humorous old-school phishing email (e.g., “Nigerian prince”). Ask: <i>“Why is this suspicious? What’s changed today?”</i>
<b>2. Mini lecture – anatomy of a phish</b>	10 min	Go over: suspicious sender domains, urgency language, fake logos, spoofed links. Introduce the idea of <i>disinformation via email</i> (e.g., false info attachments, fake petitions).
<b>3. Group work – email dissection lab</b>	15 min	In pairs or trios, teachers analyse 3 sample emails (mix of real, phishing, and gray-zone). They fill out a checklist: sender, grammar, links, emotion, disinfo claims. Share one finding per group.
<b>4. Activity – “link It back”</b>	10 min	Groups practice right-clicking/checking sender email address, hovering over links, and checking suspicious domains using tools like <a href="https://whois.domaintools.com">whois.domaintools.com</a> .

<b>5. Group discussion – What would you teach?</b>	10 min	Groups brainstorm 3 ways to teach students how to check a suspicious message or link.
<b>6. Wrap-up &amp; reflect</b>	5 min	Share one surprise + one classroom strategy from each group. Encourage future peer-sharing.

### Discussion prompts:

- What emotional tricks do these emails use to make you click?
- How could a phishing message spread disinformation?
- Why might students be especially vulnerable to link-based attacks?
- How does checking origin promote critical digital citizenship?

### Assessment Rubric

Criteria	1 – Emerging	3 – Developing	5 – Proficient
<b>Email clues identified</b>	Few, vague signs noticed	Some relevant clues named	Multiple accurate red flags with clear reasoning
<b>Tool use</b>	Struggles to trace URLs or domains	Tries tools with partial success	Confidently uses link tracing, sender ID techniques
<b>Group collaboration</b>	Minimal interaction	Some shared work	Active participation, collective conclusions
<b>Classroom application</b>	Vague ideas	1–2 classroom ideas	Clear, relevant teaching strategies for students

## 5. Recognising emotional manipulation

### Objective

Experience how AI-amplified emotional content exploits biases—and discuss how to teach this to young students.

### Materials

- Printed or digital ‘emotional content’ samples (headlines, memes, [posts](#))
- Emotional Scale Poster (low to high emotional intensity)
- [III.5.Emotional Manipulation StudentHandout.docx](#)
- Markers, sticky notes
- Access to AI-generated headlines (e.g., via ChatGPT or similar tools)

### Lesson flow (60 minutes)

Phase	Time	Activity
<b>Warm-up</b>	10 min	Display emotional headlines or memes. Ask: “What do you feel? Why?”
<b>Exploring manipulation</b>	10 min	Introduce AI-generated headlines—some neutral, some exaggerated. Groups sort by emotional intensity.
<b>Discussion</b>	10 min	Each group chooses 1-2 emotionally intense examples. Ask: “What makes this manipulative?”
<b>Classroom link</b>	10 min	In groups, plan how to guide students to recognize emotional triggers in online content.
<b>Reflection &amp; exit</b>	10 min	Each teacher shares one new strategy they’ll use with students.

### Discussion prompts

- How does emotional content spread faster online?
- Can AI amplify this effect? How?
- What age-appropriate examples can help students spot manipulation?

## Assessment rubrics

Criteria	Emerging (1)	Developing (3)	Proficient (5)
Identification of emotion	Struggles to identify emotional tone or chooses inaccurately.	Recognizes general emotion but lacks detail or confidence.	Accurately identifies targeted emotion with clear explanation.
Recognition of manipulative technique	Minimal recognition of language techniques or impact.	Some identification of exaggeration or charged words.	Clearly identifies manipulative features with examples.
Group discussion & insight	Minimal participation or unclear contributions.	Shares basic ideas or listens actively.	Engages in thoughtful discussion and builds on peer ideas.
Teaching strategy suggestion	Unclear or impractical ideas.	Basic strategies that need development.	Strong, practical, student-friendly strategy articulated.

## 6. Identifying bots

### Objectives:

- Understand how bots operate and how they spread or amplify disinformation.
- Learn practical techniques for identifying social media bots.
- Develop classroom strategies to help students recognize and question bot content.

### Materials & tools:

- Laptops or tablets with internet access
- Sample social media posts (printed or digital)
- [III.6.Bot Recognition ClassroomHandout.docx](#)
- Case Study Sheets: Real vs. Bot Accounts
- Access to tools: Botometer (<https://botometer.osome.iu.edu/>)
- Analyse real tweets and decide which ones could be bot-generated using Bot Sentinel

### Lesson flow (60 minutes):

1. Introduction & warm-up (10 minutes):
  - Ask: Have you ever suspected a post was from a bot? What gave it away?
  - Briefly explain how bots work (automated accounts, goal-driven content).
2. Case study analysis (10 minutes):
  - Pairs examine 2-3 social media profiles using printed samples or online.
  - Use the checklist to decide: bot or human?
  - Discuss indicators: posting frequency, lack of personal content, copy-paste behaviour.
3. Botometer Investigation (10 minutes):
  - Teachers explore Botometer or similar tools in small groups.
  - Enter handles (pre-approved) to analyse bot-likelihood.
  - Reflect on reliability and transparency of detection tools.
4. Impact Mapping (15 minutes):
  - Groups discuss and chart: How might bots shape student thinking?
  - Map emotional triggers, echo chambers, political targeting.
5. Create Classroom Guidelines (15 minutes):
  - Each group drafts 3 tips for students on spotting and handling bots.
  - Share with whole group for refinement.

### Discussion prompts:

- What are the telltale signs of a bot account?
- How do bots emotionally manipulate or mislead?

- What are ethical concerns in detecting and reporting bots?
- How do we teach students to avoid falling into bot-generated echo chambers?

### Assessment rubric: Spot the bot activity

Criteria	1 – Beginning	3 – Developing	5 – Proficient	Comments
<b>Bot identification</b>	Guesses without evidence	Identifies basic signs	Uses multiple indicators accurately	
<b>Critical thinking</b>	Limited reasoning	Some justification	Clear, logical justification for conclusions	
<b>Collaboration</b>	Little participation	Some group interaction	Active discussion and teamwork	
<b>Communication</b>	Unclear explanations	Some clarity	Clear, structured reasoning shared	
<b>Reflection</b>	No insight	One surface insight	Deep understanding of influence & detection	

### Wrap-up & takeaways:

- Share one classroom strategy you're excited to try.
- Reflect: What's one myth you had about bots that changed today?



## IV. Cultivating responsible Digital citizenship

**Proposed activity:** Depending on the pre-test results of students choose the most relevant activity for the class.

### 1. Think before sharing

#### Objectives:

- Understand the principles of responsible digital behaviour and information sharing.
- Build and critique a “Think Before Sharing” model to guide students.
- Practice techniques of “social correction” — constructive ways to respond to misinformation shared by peers.
- Prepare to embed these strategies into student-facing lessons.

**Duration:** 60 minutes

**Group format:** groups of 3–5

#### Materials needed:

- Flipchart paper or whiteboards, markers and sticky notes/Miro or Padlet
- [IV.1.Think\\_Before\\_Sharing\\_Handout.docx](#)
- [IV.1.Social\\_Correction\\_Role\\_Play\\_Handout.docx](#)
- Handout with real or simulated social media posts (some misleading, some neutral)

#### Lesson flow:

##### 1. Welcome & warm-up (10 min)

- Ask:  
*“Have you ever seen something online and shared it before verifying it?”*
- Show quick examples (e.g., viral image with misleading caption, fake quote).
- Short reflection on personal sharing habits.

##### 2. Mini-input: What Is responsible Digital citizenship? (10 min)

- Facilitator presents key traits: respect, accuracy, empathy, privacy, and accountability.

- Highlight research showing that social media users often shape peer behaviour.

### 3. Activity: Think before sharing model (15 min)

- In groups, teachers develop a classroom-friendly “Think Before Sharing” strategy using a simple mnemonic (e.g., **T.R.U.T.H.** = True, Reliable, Understandable, Timely, Helpful).
- They must include:
  - 5 steps/questions students ask before sharing
  - Age-appropriate language
  - 1 example post that would pass, and 1 that would fail the check

### 4. Role-play: Practicing social correction (15 min)

- Each group selects a misleading post from the handout.
- Scenario: *“Your student/friend shares this in class/group chat—how do you respond respectfully, but firmly?”*
- Groups create and act out short dialogues, demonstrating “social correction” using non-confrontational phrases like:
  - “Hey, I used to think that too—then I looked it up...”
  - “Mind if I share a different take on that?”

### 5. Share & reflect (10 min)

- Each group briefly presents their model.
- Whole-group reflection:
  - “How can we help students see themselves as responsible online citizens—not just consumers?”*
  - “What language empowers young people to self-regulate and help others?”*

### Discussion prompts:

- What are the real-world consequences of irresponsible sharing?
- How do power, peer pressure, and social status affect what students post or correct?
- How can we make digital citizenship feel *relevant* to students?

## Assessment Rubric:

Criteria	Emerging	Proficient	Advanced
Strategy clarity	Basic slogan or unclear logic	Clear 3–5 step strategy with examples	Memorable, age-adapted strategy with visuals/tools
Social correction language	Limited or judgmental phrases	Uses polite, evidence-based responses	Models empathetic and empowering dialogue
Application to classroom	Unclear or generic ideas	Some discussion of use with students	Clear integration plan for multiple grade levels
Collaboration	Uneven group input	Shared idea development	Strong team synthesis, peer coaching

## Final Summary (for all sessions)

At the end of the sequence, teachers will be able to:

1. Explain **how AI works** and its limitations.
2. **Critically evaluate** AI-generated content through ethical questioning.
3. **Detect synthetic media** using accessible tools and checklists.
4. Understand and teach emotional manipulation tactics.
5. Implement “Think Before Sharing” routines and model digital citizenship.

## 2. Netiquette use (Promoting respectful online communication)

### Objective:

- Understand key principles of netiquette (network etiquette).
- Practice applying netiquette to real classroom and disinformation scenarios.
- Reflect on how to teach students to promote respectful, thoughtful online communication.

**Duration: 60 minutes**

**Group format: Small groups of 3–4 participants**

### Materials needed:

- [IV.2.Netiquette Principles Handout.docx](#)
- [IV.2.Netiquette Scenario Cards.docx](#)
- [IV.2.Reply Rewrite Worksheet.docx](#)
- Flip chart, board/Miro, Padlet for sharing group rules
- Markers or pens

### Lesson flow:

Phase	Time	Activity
<b>1. Hook – disrespect spiral</b>	10 min	Read aloud a short fake post (e.g., “You’re just brainwashed if you believe that!”). Ask: <i>What would happen if we replied harshly?</i> Introduce “Netiquette.”
<b>2. Explore – What Is Netiquette?</b>	10 min	In pairs, teachers read through a handout listing 7 netiquette principles. They identify which ones are often forgotten or hard to apply.
<b>3. Practice – Scenario rewrite</b>	15 min	Each group draws a <b>scenario card</b> (e.g., trolling, heated comment thread, misinformation). They rewrite the comment thread applying netiquette and complete the “Reply Rewrite” worksheet.
<b>4. Gallery walk – “Before &amp; after” threads</b>	10 min	Groups post their original vs. rewritten replies. Peers walk around and vote with stickers on the most respectful and most effective rewrites.

<b>5. Discussion – Emotional triggers &amp; groupthink</b>	10 min	Reflect together: <i>Why do online discussions spiral? How can we guide students to pause, check tone, and resist online mobs?</i>
<b>6. Action plan – Class Netiquette charter</b>	10 min	Each group drafts 3 netiquette rules they would use in the classroom. One representative posts it on the flipchart.

### Discussion prompts:

- “How does tone change when we disagree online?”
- “What can students do if others escalate anger?”
- “How do trolls use emotion to manipulate?”
- “Can netiquette reduce disinformation spread?”

### Assessment rubric:

Criterion	1 – Emerging	3 – Developing	5 – Proficient
<b>Scenario analysis</b>	Vague or unclear responses	Identifies key issue, but limited rewrite	Clearly analyses tone, emotion, and risk of escalation
<b>Netiquette application</b>	Minimal or superficial rewrite	Applies 1–2 principles effectively	Applies multiple principles with nuanced tone
<b>Collaboration</b>	Limited group discussion	Some teamwork and shared ideas	Actively engaged, builds on others’ suggestions
<b>Reflection &amp; transfer</b>	Simple observations	Mentions relevance to students	Strong insights on teaching netiquette in context

### 3. Information hygiene

**Duration:** 60 minutes

**Group type:** 3–4 per group

**Objective:**

- Promote mindful content consumption.
- Encourage deliberate, responsible sharing practices.
- Equip teachers to model information hygiene for students.

**Materials**

- Projector and slides (optional)
- [IV.3.5 Point Information Hygiene Checklist.docx](#)
- [IV.3.Think\\_Before\\_You\\_Share\\_Decision\\_Tree.docx](#)
- [IV.3.Digital Citizenship Guidelines Worksheet.docx](#)
- Flipchart paper and markers/Miro, Padlet
- Coloured stickers for group voting
- Explore [News Literacy Project Tools](#)

**Lesson flow**

Phase	Time	Activity
<b>1. Warm-up discussion</b>	10 min	Think-Pair-Share: “What’s the last thing you shared online—and did you check it?” Pairs share with the group.
<b>2. Mini input</b>	10 min	Facilitator shares the definition of <i>information hygiene</i> and common types of unverified content (e.g., old news recirculated, emotional posts, satire as fact). Introduce the <i>5-Point Hygiene Checklist</i> .
<b>3. Group analysis</b>	15 min	Each group receives 2–3 short “info hygiene scenarios” (questionable WhatsApp forwards, misleading memes, etc.). They use the checklist to evaluate each item. Flipchart: Is it shareable, needs verification, or should be discarded?
<b>4. Gallery walk</b>	10 min	Groups display decisions. Teachers walk around, place coloured dots on peer posters (green = agree, yellow = unsure, red = disagree).

- |   |        |  |
|---|--------|--|
| <b>5. Discussion: social pressure &amp; sharing</b> | 10 min | Whole-group reflection on what drives irresponsible sharing (emotions, groupthink, urgency). How can we slow down? What should we model for students?                                    |
| <b>6. Wrap-up &amp; takeaway</b>                    | 10 min | Each teacher writes a personal “Information Hygiene Pledge” on a sticky note: one habit to change or adopt. Optional: create a poster titled “Think Before You Share” for classroom use. |

### Prompts for discussion

- “What makes people forward or repost false content without checking?”
- “Why is deleting something after realizing it’s false not enough?”
- “How can we model better behaviour for students?”

### Assessment rubric (Collaborative reflection)

Criteria	1 – Emerging	3 – Developing	5 – Proficient
<b>Checklist use</b>	Misses key checks or rationale unclear	Applies checklist partially with some logic	Applies full checklist with clear justifications
<b>Group collaboration</b>	One person dominates, others passive	Moderate participation across members	Equitable input, supportive team analysis
<b>Reflection depth</b>	Vague responses or blame-shifting	Some acknowledgment of habits and biases	Thoughtful, self-aware insight and actionable takeaways
<b>Scenario analysis</b>	Misclassifies most items	Classifies with some accuracy	Accurate, nuanced classification and discussion

## 4. Echo Chamber

### “Echo chamber: Seeing beyond our digital walls”

#### Objective

Understand how digital echo chambers form, recognize their influence on beliefs, and explore ways to teach students critical content navigation across perspectives.

**Duration:** 60 minutes

#### Materials needed

- Chart paper or digital collaborative board (e.g., Jamboard, Miro)
- [IV.4.Signs\\_of\\_an\\_Echo\\_Chamber.docx](#)
- [IV.4.Echo\\_Chamber\\_Simulation\\_Cards.docx](#)
- Coloured pens or highlighters
- Watch [TED Ed: How social media filters your worldview](#)

#### Lesson flow

Phase	Time	Activity
<b>1. Warm-up scenario</b>	10 min	Read aloud a fictional but realistic student statement: <i>“Everyone agrees this is true, it’s all over my feed.”</i> Ask: “Why do you think they believe this?” Brainstorm first ideas.
<b>2. Echo chamber simulation</b>	20 min	In groups of 4, teachers draw 6–8 “info cards” representing statements (some reinforcing, some dissenting). They sort what they’d “share,” “ignore,” and “disagree with.” Gradually remove dissenting ones. Reflect on how it shapes perception.
<b>3. Mapping groupthink</b>	10 min	On chart paper, each group maps how their selection affected their collective beliefs. Compare across groups: did any end up seeing mostly one perspective?
<b>4. Strategy brainstorm</b>	15 min	Introduce “Signs of an Echo Chamber” handout. Groups list classroom strategies to: <ul style="list-style-type: none"> <li>• Identify narrow feeds</li> <li>• Teach respectful cross-viewpoint engagement</li> <li>• Prevent algorithmic isolation</li> </ul>
<b>5. Wrap-up &amp; debrief</b>	5 min	Round-robin: Each group shares one takeaway & one idea to test in their classroom. Encourage reflection on digital openness.



## Discussion prompts

- “What types of content got filtered out in your group?”
- “How can algorithms reinforce these bubbles?”
- “How can we model ‘perspective broadening’ with students?”
- What all-world consequences can arise from echo chambers?

## Assessment rubric

Criteria	1 – Emerging	3 – Developing	5 – Proficient
<b>Understanding of Echo chambers</b>	Limited recognition	Can define & give 1–2 examples	Clear grasp, with classroom implications
<b>Collaboration in simulation</b>	Passive or disengaged	Some contribution	Fully engaged with active role-taking
<b>Strategy development</b>	Generic suggestions	Some targeted ideas	Practical, age-appropriate classroom applications
<b>Reflection quality</b>	Vague or off-topic	Identifies 1 key insight	Thoughtful link between activity & teaching practice

## 5. Understanding & managing your Digital footprint

**Duration:** 60 minutes

### Objective:

- Explore how personal data is collected across platforms.
- Identify the long-term implications of digital footprints.
- Practice techniques to manage digital identity responsibly.

### Materials & tools:

- Projector or screen with internet access
- Access to online footprint analysis tools (e.g., [Trace My Shadow](#), [Google Activity](#))
- Chart paper / sticky notes / markers/Padlet, Miro
- [IV.5.Footprint Check Student Exercise.docx](#)
- [IV.5.Footprint Mapping Template.docx](#)
- **Interactive Tool (10 min):** Visit <https://www.digitalfootprintcheck.com> (demo site)
- **Group simulation (10 min):** Create a fictional character and chart their digital trail over a week.
- **Discussion (5 min):** How can students protect their identity?

### Flow of activities:

Phase	Time	Activity
<b>1. Introduction</b>	10 min	Icebreaker: "What's the last digital trace you left?" Discussion: What is a digital footprint? Visible vs invisible data traces.
<b>2. Personal mapping</b>	10 min	In pairs, teachers list all platforms/services they use. Using <i>Footprint Mapping Template</i> , write down what data they think each collects.
<b>3. Tool walkthrough</b>	10 min	Small groups try out <i>Trace My Shadow</i> or check their <i>Google Activity</i> . Add surprises or new findings to their map.
<b>4. Data chain simulation</b>	10 min	Each group gets a scenario (e.g., sharing location, signing up for a quiz app). They trace how data might travel between platforms/companies.

<b>5. Strategy brainstorm</b>	10 min	Discuss strategies to teach students: privacy settings, limiting oversharing, digital reflection.
<b>6. Wrap-up &amp; commitments</b>	10 min	Share one digital habit each teacher will change or model in class. Distribute <i>Footprint Check</i> handout to adapt for students.

### Discussion prompts:

- “How much control do we have over our footprint?”
- “Which data traces are intentional vs passive?”
- “How would you explain this to 10-year-olds? To teens?”
- “Should we ever ‘scare’ students about their footprint—or just empower them?”

### Assessment rubric for participation

Criteria	1 – Emerging	3 – Developing	5 – Proficient
<b>Footprint awareness</b>	Lists few or vague data points	Identifies several platform-specific data points	Comprehensive and reflective mapping with insights
<b>Tool exploration</b>	Minimal engagement with tools	Uses tools, notes key data	Deep exploration, compares tools, shares surprising finds
<b>Scenario application</b>	Misses steps in data journey	Identifies basic data flow	Clearly maps complex data chain implications
<b>Collaboration</b>	Limited input or engagement	Contributes ideas, participates in discussion	Facilitates discussion, helps others reflect

## 6. The disinformation web

### Objective:

To simulate how disinformation spreads and discover how coordinated action between schools, parents, local media, and community partners can slow or stop its spread.

**Duration:** 30–40 minutes

Materials needed:

- Index cards or slips of paper
- Markers or pens
- Yarn or string (for the web simulation)
- Flipchart or whiteboard

### Step-by-step instructions:

#### 1. Introduction (5 minutes):

Briefly present a *fictional disinformation scenario*. Example:

A rumour has started online that your school is installing hidden cameras in student lockers. It's being shared in parent WhatsApp groups and a local blog.

Explain that participants will play different roles in a simulated community affected by the rumour.

#### 2. Role assignment (5 minutes):

Randomly assign roles to participants, such as:

- Teacher
- Principal
- Parent
- Student
- Local journalist
- IT coordinator
- Community activist
- School board member

Give each participant a card with their role and some basic information (e.g., "You just heard the rumour from a parent").

### 3. Simulation: Spread the rumour (10 minutes):

Use yarn to visually represent the flow of information: each participant who hears the rumour passes the yarn to the next person they tell. Very quickly, a web forms.

Pause when the rumour has spread significantly. Ask:

- “How could this have been stopped earlier?”
- “Who could have intervened?”
- “What were the points of trust or authority?”

### 4. Group reflection & debrief (15 minutes):

As a group, answer:

- What *systems* (e.g., media literacy, school-parent communication) could have helped?
- How might better relationships with the community (e.g., local media, PTA, youth centres) reduce harm?
- What’s the school’s role *beyond the classroom* in combating disinformation?

Write reflections on the board. Identify 3–4 **strategic actions** schools can take to build stronger community partnerships.

### Teacher collaboration & disinformation response rubric

Criteria	Excellent (4)	Good (3)	Satisfactory (2)	Needs Improvement (1)
<b>Understanding of disinformation</b>	Demonstrates thorough understanding of disinformation types, sources, and impacts.	Understands most disinformation forms and general implications.	Shows basic awareness, but some misconceptions persist.	Lacks clear understanding or confuses terms like misinformation, disinformation, and malinformation.
<b>Community engagement</b>	Proactively collaborates with parents, local media, and NGOs; initiates community projects.	Participates in planned school-community activities and encourages collaboration.	Supports collaboration when guided but does not take initiative.	Avoids or overlooks community-based collaboration.

<b>Classroom integration</b>	Seamlessly integrates digital/media literacy into lessons with real-world relevance.	Occasionally integrates media literacy or disinformation awareness in teaching.	Rarely connects curriculum to disinformation or digital risks.	No evidence of integration.
<b>Use of school policy</b>	Actively shapes or promotes policy and guidelines for responding to disinformation.	Understands and adheres to school policy and digital behaviour expectations.	Aware of school policy but applies it inconsistently.	Unaware or disregards policies related to disinformation.
<b>Collaboration with colleagues</b>	Leads or mentors staff in community-based digital literacy initiatives.	Collaborates with colleagues in training or workshops.	Participates occasionally in joint efforts.	Works in isolation with no peer collaboration.
<b>Ethical digital behaviour modelling</b>	Consistently models ethical behaviour and teaches verification techniques.	Often demonstrates responsible use of digital tools.	Demonstrates ethical behaviour but lacks emphasis in teaching.	Rarely demonstrates or discusses digital ethics.

### Scoring guide:

- **21–24:** Exemplary – Teacher is a leader in combating disinformation through school-community collaboration.
- **16–20:** Proficient – Teacher applies knowledge and promotes responsible practices with moderate initiative.
- **10–15:** Developing – Teacher shows some effort but needs support to deepen understanding and collaboration.
- **Below 10:** Beginning – Teacher requires significant development in both understanding and practice.

### Takeaway:

Wrap up by emphasizing:

“Fighting disinformation is not just about detecting lies—it’s about building trust and clear communication. Collaboration between schools and the community turns a rumour into a teachable moment instead of a crisis.”

# Adapting activities for adolescents: A teacher's guide

## 1. Define clear learning goals

Before adaptation, clarify what you want students to:

- **Know** (e.g., "What is a deepfake?")
- **Do** (e.g., "Verify the source of an image")
- **Feel** (e.g., "Be confident in questioning online content")

Use age-appropriate language in objectives:

Instead of: "Understand AI detection mechanisms"

Try: "Learn how to spot a fake photo made by a computer."

## 2. Use youth-relevant examples

Replace adult scenarios with issues closer to adolescent experiences:

Original scenario	Youth-friendly version
AI in grading essays	AI writing tools and homework apps
Deepfakes of politicians	TikToks of influencers saying strange things
Disinformation in elections	Rumours about school events or celebrities
Incorporate platforms they use (e.g., TikTok, Instagram, Discord) to make it real.	

## 3. Make it interactive & visual

Adolescents respond well to visual, creative, and game-based learning.

### Suggestions:

- Use memes, GIFs, and videos to explore misinformation.
- Turn identification of fake news into a quiz or "spot the fake" game.
- Encourage students to create **infographics, posters, or videos** to teach peers.

#### 4. Facilitate critical dialogue

Use peer discussion methods to make abstract ethical issues tangible:

- **Ethics debate carousel** → Adapt with questions like:  
“Should students be allowed to use AI in art class?”
- **Four corners activity** → Read a statement (e.g., “I trust AI more than people online”) and let students move to corners to show agreement/disagreement, then explain why.

#### 5. Introduce easy-to-use tools

Use **student-accessible tools** to demonstrate key concepts:

- **AI image detection** → Show how to reverse image search using Google Lens
- **Bot detection** → Use simulations or websites like [botometer.org](https://botometer.org)
- **Media verification** → Let students use browser extensions like InVID or use fact-checking websites

#### 6. Focus on empowerment, not fear

Teens may feel overwhelmed by AI or disinformation. Frame lessons around **digital resilience** and **critical thinking**:

- Emphasize that they **can** spot and report manipulation.
- Discuss how **trust, empathy, and responsible sharing** are superpowers online.

#### 7. Scaffold the complexity

Break down big ideas into layers:

- Start with the concept (e.g., “What’s fake online?”)
- Then the method (e.g., “How can I tell?”)
- Then the action (e.g., “What should I do about it?”)

Use tiered worksheets, levelled questions, and graphic organizers to support comprehension.

#### 8. Encourage student-led activities

Let adolescents take ownership:

- Form “digital detectives” teams to monitor disinformation in their school communities.
- Assign mini-projects like “Create a fake story and then debunk it.”
- Host classroom “press conferences” to simulate crisis response.



## 9. Integrate with curriculum subjects

Link activities to subject areas:

Subject	Integration idea
Language arts	Analyse persuasive techniques in disinformation
Civics	Explore how false information affects democracy
ICT	Examine how AI tools work and their limitations
Visual arts	Study visual manipulation and deepfakes

## 10. Include self-reflection

Use simple self-checks like:

- “One thing I believed before, but now I question is...”
- “What would I tell a younger student about AI and fake news?”

## KEY TOOLS+RESOURCES

### Fact-checking websites:

- [FactCheck.org](#), [Snopes](#), [PolitiFact](#), [Originality.AI](#)

### Media literacy resources:

- [MediaWise](#), [News Literacy Proje](#)

### AI verification tools:

- [OpenAIDetector](#), [Detecting.AI](#), [GoogleReverselImageSearch](#)

***Commit some time to explore “specific detection tools!”***

# APPENDIX

## Key terms and definitions

### 1. Digital literacy

*The ability to use digital tools and technologies effectively, safely, and critically.*

This includes skills such as evaluating online information, navigating websites, using apps, and understanding digital rights and responsibilities.

### 2. Cybersecurity

*The practice of protecting digital systems, networks, and data from unauthorized access, attacks, or damage.*

It involves using secure passwords, recognizing online threats, and following safe online behaviours.

### 3. Resilience building

*The process of developing the ability to recover from setbacks, adapt to change, and keep going in difficult situations.*

In education, it helps students manage stress, overcome failure, and grow stronger emotionally and socially.

### 4. Problem solving

*The ability to identify a challenge, analyse it, and find effective solutions through reasoning and creativity.*

It involves critical thinking, decision-making, and applying knowledge in new ways.

### 5. Critical thinking

*The skill of thinking clearly and logically, questioning assumptions, evaluating evidence, and forming reasoned conclusions.*

Essential for making informed decisions and resisting misinformation.

### 6. Industry relations

*The connections and collaborations between schools or educational institutions and the business or professional sectors.*

These relationships help align learning with real-world skills and career opportunities.

## 7. Community relations

*The ongoing engagement between schools and the local community, including families, organizations, and stakeholders.*

Strong community ties support inclusive education, resource sharing, and social responsibility.

## 8. Vulnerable students

*Learners who face additional challenges that may hinder their educational progress or well-being.*

This can include students affected by poverty, disability, trauma, language barriers, or discrimination.

## 9. Intersectionality

*A framework for understanding how different aspects of a person's identity (e.g., race, gender, class, disability) interact and can lead to overlapping forms of disadvantage or discrimination.*

It promotes equity by recognizing diverse experiences and needs.

## 10. Bullying

*Repeated, intentional behaviour that hurts, threatens, or humiliates another person, often involving a power imbalance.*

Bullying can be physical, verbal, relational, or online (cyberbullying), and it seriously affects emotional and social development.

## 11. Disinformation

*False information deliberately created and spread to deceive or mislead people.*

Example: Fake news articles designed to manipulate public opinion.

## 12. Misinformation

*False or inaccurate information shared without intent to mislead.*

Example: Sharing a wrongly captioned image believing it to be true.

## 13. Malinformation

*Genuine information used to harm a person, organization, or country.*

Example: Leaking private data to embarrass someone.

#### 14. Fake news

*Fabricated news stories presented as legitimate journalism, often shared on social media.*

Example: Viral headlines claiming false scientific discoveries.

#### 15. Deepfake

*Synthetic media (usually video or audio) generated by AI to imitate real people, often convincingly.*

Example: A video showing a politician saying things they never actually said.

#### 16. Algorithmic amplification

*The way social media platforms use algorithms to boost the visibility of content—sometimes unintentionally promoting disinformation.*

Example: Sensational or misleading posts going viral due to engagement-focused algorithms.

#### 17. Echo chamber

*An online environment where people are only exposed to opinions that reinforce their own, limiting critical thinking.*

Example: A social media feed curated to only show one political viewpoint.

#### 18. Confirmation bias

*The tendency to believe or share information that confirms one's existing beliefs, regardless of its truth.*

Example: Ignoring credible sources that contradict a preferred opinion.

#### 19. Media literacy

*The ability to critically analyse and evaluate media content, sources, and intentions.*

Key to helping students and adults spot and question disinformation.

#### 20. Fact-checking

*The process of verifying information using reliable sources before sharing or publishing.*

Organizations like Snopes or PolitiFact help verify viral claims.

## **21. Source credibility**

*Judging how trustworthy a source is based on its history, expertise, and transparency.*

Example: Prioritizing information from peer-reviewed journals over anonymous blogs.

## **22. Digital footprint**

*All the data a person leaves behind online, including shared content, comments, and posts.*

Being aware of what we share helps limit the spread of harmful content.

## **23. Bot**

*An automated software program that can post or share content on social media, often used to spread disinformation quickly.*

Example: Thousands of bots retweeting a false story to make it trend.

## **24. Critical thinking**

*The skill of logically evaluating arguments, evidence, and claims.*

Essential for resisting manipulation and forming balanced opinions.

## **25. Prebunking (Inoculation theory)**

*A proactive strategy that exposes people to weakened forms of false claims in advance, to help them resist persuasion.*

Used in educational settings to build “mental immunity” to disinformation.

## **26. Debunking**

*The act of exposing and correcting false information after it has been circulated.*

Effective when done clearly, respectfully, and with solid evidence.

## **27. Information hygiene**

*Personal habits that reduce the risk of spreading false information, such as checking sources and avoiding emotional sharing.*

Like washing hands to prevent illness—think before you share.

## **28. Cybersecurity awareness**

*Understanding digital threats—including disinformation—as part of a broader effort to keep individuals and institutions safe online.*

Includes spotting phishing, manipulation, and media forgery.

## **29. Narrative manipulation**

*Strategic shaping of public perception through coordinated, misleading storytelling.*

Often used in political or ideological campaigns.

## **30. Platform responsibility**

*The ethical and legal obligations of social media companies to limit harmful content and ensure transparency.*

Examples: Labelling altered media, removing bots, adjusting recommendation algorithms.

## **31. Phishing**

*Phishing is a form of online deception in which attackers trick individuals into revealing sensitive personal information—such as passwords, credit card numbers, or login details—by pretending to be a trustworthy source through emails, messages, or fake websites.*

Examples: Sending text from misleading email asking for information.

## **32. Social engineering**

*Social engineering is the act of manipulating people into giving up confidential information or performing actions that compromise security. It relies on psychological manipulation, not technical hacking, and often involves building trust, creating urgency, or impersonating authority to deceive the target.*

Examples: Calling someone on phone manipulating the person to send money for a loved one in danger.