Co-funded by the European Union

ACTIVATING LESSON SCENARIO
developed as part of the project

‘INNOVATIONS IN SCHOOL EDUCATION’

TOPIC

“Deepfake – Can We Trust What We See and Hear?” (Visual and Audio Manipulation – Risks and Methods of Detection)

1. Lesson objectives

The student:

- understands what deepfake is and how it is created,
- knows the most common uses and risks related to deepfake technology,
- can identify features of a fake video or audio recording,
- understands how deepfake can be used in disinformation and cyberbullying,
- develops critical skills for receiving multimedia content online.

2. Target group

Primary school students

3. Teaching methods

- Brainstorming
- Mini-lecture with multimedia presentation
- Analysis of short materials (screenshots/videos)

- Group exercise – “Real or Manipulated?”
- Guided discussion
- Individual reflection

4. Teaching aids / sources

- Computer, projector, interactive whiteboard
- Sample screenshots or short, fictional (safe) videos and audio recordings that have been manipulated (without using real people)
- “10 Deepfake Warning Signs” card
- List of educational and fact-checking websites:
- EU: https://edmo.eu, https://euvsdisinfo.eu
- Poland: https://demagog.org.pl
- Czech Republic: https://manipulatori.cz
- Slovakia: https://infosecurity.sk
- Ukraine: https://www.stopfake.org

5. Lesson procedure (45 min)

1. Introduction – Do recordings always show the truth? (5–7 min)

1. Brainstorming – students’ experiences (2–3 min)

- The teacher asks students to give examples of recordings from the internet that looked:
- “strange”, “unnatural”,
- “too shocking to be real”,
- raised doubts about whether they really showed a given situation or person.

Answers may be written on the board in two columns:
“Recordings I trust” / “Recordings that seemed suspicious.”

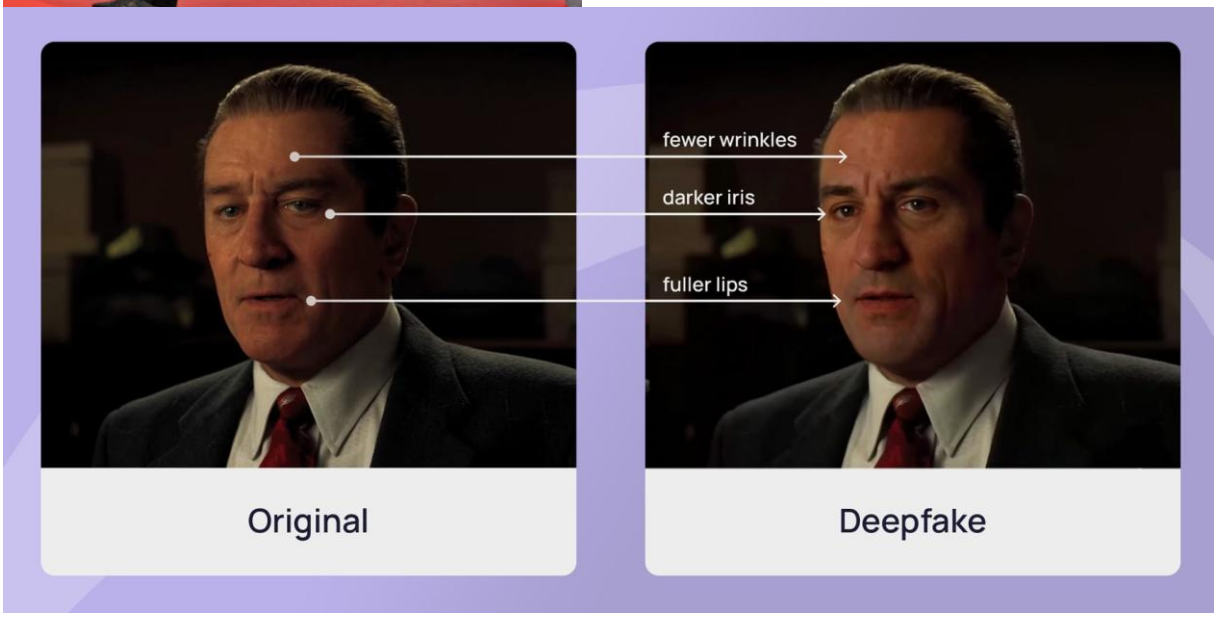
The goal is to show that not all videos and recordings online are reliable, even if they look professional.

2. Guiding questions for a short discussion (2 min)

- Can we always believe what we see in a video or hear in a recording?
- Can technology make someone say or do something they never actually said or did?
- Why is it becoming harder to distinguish real recordings from fake ones?
- Have you ever seen a video that looked like a joke or was altered by filters and apps? How does that affect belief in its authenticity?

3. Short examples for analysis (optional, 1–2 min)
The teacher may show 2 images or short slides (fictional):

- a photo/video showing a well-known person in an unusual situation (e.g. speaking in a foreign language),
- a fragment of a manipulated recording with clearly unnatural lip movement or sound.



Example 1 – Famous person in an unusual, unexpected situation
An image shows a celebrity (e.g. on the red carpet) in an unusual moment – such as tripping,

falling, or another surprising situation that would be hard to see outside of context. This is a perfect starting point to discuss how easily context can be manipulated.

Example 2 – Deepfake patterns – disturbing facial details
An image compares natural and manipulated video: visible differences around the mouth, gaze, facial proportions. Such visual examples make it easier to spot subtle signs of forgery.

Then ask:

- Does this photo look real? What about it raises doubts?
- How can we know whether the material is trustworthy?

4. Teacher's commentary – factual information (2 min)

- A deepfake is a fake video, image, or sound created with artificial intelligence, pretending to be genuine material.
- This technology can replace a person's face, voice, or even entire behaviours in a recording so that it looks extremely realistic.

Purposes of using deepfake:

- entertainment (e.g. filters, humorous videos),
- fraud and money scams, impersonating famous people,
- blackmail, cyberbullying, reputation damage,
- political and social disinformation, creating false evidence.

The risk: in the deepfake era, it is becoming harder to distinguish truth from manipulation, so we must carefully analyse recordings before trusting or sharing them.

2. Mini-lecture: What is deepfake and what risks does it bring? (10–12 min)

1. Introduction (1 min)

The teacher asks:

Do you believe everything you see in online videos?

Does a video recording always show the truth?

Emphasises that in the era of new technologies, "seeing" does not always mean "believing," because artificial intelligence can create fake recordings that look real.

2. Definition of deepfake (2 min)

Deepfake is a technology based on artificial intelligence, where algorithms (neural networks) analyse hundreds or even thousands of photos and recordings of a real person to generate a fake image, video, or voice that looks and sounds like the original.

The name comes from:

deep learning – the technology behind the phenomenon,

fake – meaning false.

Purpose: to create a recording so realistic that it is difficult to distinguish from the truth.

3. How is a deepfake created? (2–3 min)

The algorithm:

Collects data – photos, videos, and voice samples of the chosen person (often public materials from the internet).

Learns the person's facial expressions, lip movements, and tone of voice to reproduce them.

Generates a new video or audio in which the person appears to say or do something they never actually said or did.

Example: creating a recording in which a celebrity advertises a product they never used, or a politician says something controversial they never actually said.

4. Examples of use (2 min)

Positive / neutral uses:

- Entertainment – TikTok filters, face swapping in films, dubbing in various languages.
- Special effects in cinema (e.g. "de-aging" actors in films).

Negative / harmful uses:

- Political disinformation: fake speeches that may change voters' opinions or destabilise social situations.
- Impersonation of famous people: celebrities, politicians, or even friends – to extort money.

- Cyberbullying and blackmail: creating false compromising materials to shame or intimidate someone.
- Financial crimes: e.g. fake voice calls from a company director ordering a bank transfer.

5. Risks of deepfake for society (2 min)

- Threat to democracy: fake recordings can influence election results or cause unrest.
- Decline of trust in media: it is becoming harder to tell truth from falsehood – people may start doubting even real recordings.
- Reputation damage: fabricated materials can ruin the personal and professional life of innocent people.
- National and corporate security: deepfakes can be used for fraud, espionage, cyberattacks.
- Psychological effects for victims: shame, stress, depression, fear of further attacks.

(The teacher may add a short fictional story, e.g.: "In one country, a video was spread showing the prime minister announcing surrender during a conflict – the recording was fake, but it caused panic among citizens.")

6. Warning signs of deepfake (3–4 min)
The teacher presents the "10 Deepfake Warning Signs" card, discussing each with examples:

- Unnatural lip movements – words don't match lip motion.
- Uneven lighting on the face, strange shadows, or missing reflections in the eyes.
- Distorted or blurred areas of the image, especially during head movements.
- Eyes that do not blink naturally.
- Delayed sound compared to lip movement.
- Metallic tone of voice, distortions, "cuts" mid-sentence.
- No source in credible media – recording appears only in unknown channels.
- Content too shocking to be real – e.g. a politician insulting an entire country, a celebrity admitting to illegal acts.

- The recording spreads quickly but lacks other confirming materials.
- No additional evidence (e.g. press releases, eyewitness accounts).

7. Summary (1 min)

- Deepfake is not just technological fun – it is also a threat to truth, people's reputations, and societal security.
- That's why we should always check the source of a recording, never trust material just because "we see it," and use multimedia verification tools (fact-checking, reverse image search).
- The teacher announces that in the next part of the lesson, students will have the opportunity to test their skills in recognising fake recordings.

3. Group exercise – "Real or Manipulated?" (15–20 min)

1. Group division and preparation of materials (2 min)

The class is divided into 3–5 groups of 3–4 students each.

Each group receives a set of 2 short recording descriptions:

1 real recording – authentic, confirmed events or fictional materials created as "credible" – https://www.youtube.com/watch?v=bquB_pKPlYk

1 manipulated recording (deepfake) – created for educational purposes, with visible manipulation features (e.g. strange lip movements, no source, synthetic voice) – https://www.youtube.com/watch?v=cQ54GDm1eL0

Each group also receives the "10 Deepfake Warning Signs" card as an analysis aid.

10 Deepfake Warning Signs
(can be printed in A4 format and handed out to each group)

10 signs that a recording may be a deepfake:

- Unnatural lip movement – no full sync with sound.
- Unrealistic eye blinking – too rare or unnatural.
- Strange shadows and lighting – light on the face doesn't match the background.

- Blurring or image artefacts – especially around the mouth and eyes.
- Change in voice tone – voice sounds “metallic” or synthetic.
- Lack of smooth motion – sudden jumps or “sliding” image elements.
- No source or unknown source of the recording – no logos, date, or author info.
- Face looks too perfect or unnatural – no wrinkles, overly smooth skin.
- Body language mismatch – gestures or facial expressions inconsistent with speech content.
- Time pressure or sensational context – the recording appears suddenly in a crisis situation.

2. Group task (12–14 min)

Step 1 – Recording analysis (5–6 min)
Students jointly watch or analyse each recording/image:

Identify suspicious elements:

- unnatural lip or eye movements,
- lack of source or news channel logo,
- image distortions,
- unusual voice tone or sound delay,
- content too sensational to be true.
  (Students may mark these elements with a coloured marker on printouts or write them down on the analysis sheet.)

Step 2 – Authenticity assessment (3–4 min)

The group decides whether the recording is:

- real,
- manipulated (deepfake),
- hard to assess (if uncertain).

They give a short justification, using the warning signs.

Step 3 – Determining the possible purpose of creating the recording (3–4 min)

Students consider:

Why might someone have created such material?

Could the purpose be:

- to evoke emotions (fear, anger, laughter),
- political manipulation,
- blackmail, mocking a specific person,
- financial gain (e.g. ads, scams),
- humorous or entertainment purposes.

3. Analysis table (to be completed in groups)

Recording (no.)	Real or fake?	What warning signs did you notice?	Possible purpose of creating the deepfake	Where could the recording be verified?
1				
2				

4. Presentation of results (3–5 min)

Each group discusses their recording and presents:

- their decision (real/fake),
- main warning signs they noticed,
- possible purpose of creating the recording,
- where the authenticity could be checked (e.g. Demagog.org.pl, StopFake.org, reverse image search).

The teacher adds missing elements, pointing out additional verification methods.

5. Summary of the exercise (1–2 min)

- Not every recording on the internet is real, and a deepfake can look very convincing.
- To avoid being deceived:

Analyse image and sound details,

Check sources and fact-checking tools,

Don’t trust recordings just because they look real.

4. Discussion: Can we trust what we see on the internet? (8–10 min)

1. Purpose of the discussion

- Help students understand the impact of deepfake on the credibility of information online.
- Develop critical skills in analysing video and audio, even if they look realistic.
- Work together to create caution rules for using online visual materials.

2. Questions for students *(for circle discussion or “snowball” method)*

1. Is video recording always proof of truth?

What situations might make a video misleading, even if it looks real?

Have you ever seen online content that turned out to be edited or a joke?

2. How does deepfake affect our trust in information?

Does awareness of deepfakes make it harder for us to believe real news?

Could this lead to a “I don’t believe anything I see online” mindset?

3. Why do fake recordings spread faster than corrections?

Do emotions and sensationalism make us click “share” faster?

Why do corrections or fact confirmations reach fewer people?

4. How can we protect ourselves from deepfake scams?

What should be our first steps when we see a shocking recording?

Should we always trust recordings sent by friends?

What tools or sites can help check authenticity (e.g. Demagog.org.pl, StopFake.org, reverse image search)?

5. Will it be even harder in the future to distinguish truth from visual manipulation?

How might AI development affect our ability to detect fakes?

Will fact-checking technology keep up with deepfake growth?

3. Discussion techniques

"Raise your hand – two sides": the teacher asks "Can video be proof of truth?" – students choose "yes" or "no" sides and argue their position.

Mind map on the board: centre topic "Can we trust recordings online?", around it arguments "for" and "against".

Reference to earlier exercises: teacher recalls group examples – which features made the recordings suspicious?

4. Teacher's conclusions (summary 2–3 min)

Deepfake is a powerful manipulation tool that can change how we perceive reality.

Even real recordings can be taken out of context or used misleadingly – caution is necessary always, not just with deepfakes.

Golden caution rule:

We do not believe a recording just because "we see it with our own eyes."

We always look for additional sources and confirmations (media, official statements, fact-checking).

In the age of AI, critical thinking is more important than ever – technology can fool the eyes and ears but cannot replace a conscious, analytical audience.

By sharing unchecked recordings, we ourselves can unintentionally become part of the disinformation problem.

Co-funded by
the European Union

5. Summary and reflection (7 min)

Students finish the sentences:

"I understood that deepfake…"

"The most suspicious thing in fake recordings is…"

"Before sharing a video or recording, I will check…"

"The most dangerous thing about deepfake is that it can…"

Joint class list: "5 ways to detect deepfake", e.g.:

I analyse image details – eyes, mouth, shadows.

I check the recording source in credible media.

I look for the same information in several independent sources.

I don't trust recordings that evoke extreme emotions and have no confirmation.

I use multimedia verification tools (e.g. reverse image search).

6. Glossary

Term	Definition
Deepfake	Fake video or audio created using artificial intelligence, looking like real material.
Visual manipulation	Deliberately altering an image to falsify reality.
Disinformation	Spreading false content to mislead audiences.
Content verification	Checking the truthfulness of materials, sources, and context in multiple places.
Warning signs	Characteristic elements indicating that a video or audio may be manipulated.

Co-funded by
the European Union

7. Methodological guide for the teacher

1. Preparation for the lesson

Selection of materials:

Use fictional or educational examples of deepfakes to avoid controversy, infringing on real people's rights, or causing unnecessary emotions.

Avoid political, religious, violent, or compromising content – materials should be neutral and safe.

Ensure diversity of formats – short video clips, audio recordings, screenshots, situation descriptions – to show that deepfake can take various forms.

Technical preparation:

Check equipment (projector, speakers) and materials before the lesson to avoid technical problems.

Make sure all video/audio files are saved offline to avoid using potentially unsafe online sources during the lesson.

2. Conducting the lesson

Introduction:

Start with examples familiar to students (app filters, video memes, voice edits) – this helps them understand what deepfake is and how easily a recording can be altered.

Encourage sharing personal experiences, but ensure no one feels judged for their answers.

Mini-lecture:

Use simple, clear language when explaining technical terms (e.g. "neural networks" → "a computer program that learns from many pictures how a person's face looks").

Insert short questions to students during the lecture to keep their attention and engagement.

Exercises:

Ensure each group has at least one student who can operate a computer or interactive whiteboard if analysing recordings.

Provide students with analysis tools ("deepfake warning signs" card, checklist questions) to make group work easier and more structured.

After the exercises, emphasise that difficulty in recognising a fake recording is normal – even experts can be wrong.

3. Moderating the discussion

Use open-ended questions: "What do you think?", "Why do you think so?", "Does anyone have a different opinion?"

Do not judge answers as "right" or "wrong" – show that every statement can be a starting point for analysis.

If a real-life example is brought up in class, make sure it is presented neutrally and does not harm third parties.

Emphasise that deepfake is technology, not the victim's fault – anyone can be deceived, which is why it is worth learning recognition methods.

4. Safe atmosphere

Establish the rule: "We don't mock others' opinions, we don't judge those who have ever been deceived."

If a student shares a personal experience (e.g. clicking on a fake video), thank them for their courage instead of criticising their behaviour.

Ensure analysed materials are not drastic, aggressive, or compromising, even fictionally – the aim is to learn analysis, not to cause unpleasant emotions.

5. Educational goal and key lesson messages

Critical thinking: Students should understand that a recording is not always proof of truth, and technology can realistically manipulate image and sound.

Caution online: Sharing unchecked materials can contribute to spreading disinformation and harm others.

Conscious media use: It is worth verifying content in multiple sources, using fact-checking websites, and seeking confirmation in credible media.

Empathy: Deepfake is often a tool for cyberbullying – we should respond when someone is a victim of fake recordings.

Co-funded by the European Union

8. Lesson extensions

Homework: Students search for information on cases of deepfake use in the world (e.g. in politics, advertising, film) and assess their impact on audiences.

Class poster: "5 rules for caution with online recordings" – created jointly to reinforce knowledge.

Mini-project: Prepare a short educational campaign for other students in the school – how to recognise deepfake and avoid being fooled.

9. Scientific and educational sources

EDMO – European Digital Media Observatory
https://edmo.eu
– European knowledge hub on disinformation, contains reports and analyses on new forms of visual manipulation, including deepfakes.

EUvsDisinfo – project of the European External Action Service
https://euvsdisinfo.eu
– Database of analysed disinformation cases, including fabricated videos, educational tools on recognising false content.

EDMO Taskforce on AI and Disinformation (2023)
https://edmo.eu/research/ai-disinformation-taskforce
– EU reports and recommendations on threats from the use of artificial intelligence in creating fake recordings.

UNESCO – Media and Information Literacy Curriculum for Teachers
https://unesdoc.unesco.org/ark:/48223/pf0000192971
– Educational programme for teachers on critical thinking and media analysis, including audiovisual manipulation.

Polish fact-checking sources
Demagog.org.pl – Poland's largest fact-checking portal
https://demagog.org.pl
– Analyses of false content, including examples of fabricated recordings and multimedia verification guides.

Czech and Slovak sources
Manipulátoři.cz
https://manipulatori.cz

– Czech educational and analytical portal specialising in tracking manipulations, including deepfakes used in disinformation campaigns.

Demagog.cz
https://demagog.cz
– Czech equivalent of Poland's Demagog, includes analyses of video recordings used to manipulate public opinion.

Demagog.sk
https://demagog.sk
– Slovak fact-checking portal, educational materials on fake news and new forms of digital manipulation.

Infosecurity.sk – Inštitút pre bezpečnostnú politiku
https://infosecurity.sk
– Analyses of information threats, reports on video disinformation and fake recordings in the Central and Eastern European region.

Ukrainian sources
StopFake.org
https://www.stopfake.org
– Ukrainian portal fighting disinformation, including analysis of fabricated recordings, fake war videos, and online audiovisual manipulations.