



AI Guide

ONLINE COURSE CONTENT

MODULE 2: FACT-CHECKING AI-GENERATED INFORMATION

2025-1-LT01-KA220-HED-000354784

Artificial Intelligence for Green Understanding, Insights & Decision Empowerment



CAM CONSULTING



SQUAREDEV



Co-funded by
the European Union

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the National Agency of Erasmus+ Programme in Lithuania – Education Exchanges Support Foundation (ŠMPF). Neither the European Union nor the ŠMPF can be held responsible for them.

Table of Contents

Fact-checking AI-generated information	3
A. General Information	3
B. Content for STEM HE Teachers.....	3
Highlights.....	3
Introduction.....	3
<i>Why AI-Generated Information Cannot Be Automatically Trusted.....</i>	<i>3</i>
<i>Understanding Fact-checking</i>	<i>5</i>
<i>Types of AI-Generated Information That Require Verification</i>	<i>5</i>
Low-risk information	6
Medium-risk information	6
High-risk information.....	6
<i>Common Types of AI Errors</i>	<i>7</i>
Hallucinated facts.....	7
Fabricated citations.....	7
Incorrect numerical information.....	7
False quotations.....	7
Outdated information.....	7
Missing context	7
<i>Frameworks for Evaluating AI-Generated Information</i>	<i>8</i>
The SIFT Method	8
The CRAAP Test	8
References	9
Green Decision Context	9
Videos/Interactive Resources	10



Fact-checking AI-generated information

A. General Information

Module No	Module2
Subtopic	Fact-checking AI-generated information
AI Literacy Component	Misinformation detection, hallucination awareness, source triangulation, and scientific evidence verification.
Decision Making Skills	Evidence verification, rejecting unsupported claims, and validating information.
Learning Objectives for STEM HE Teachers	Apply fact-checking procedures to AI-generated claims using reliable scientific and technical sources.
Learning Objectives for STEM Students	Verify AI-generated environmental claims using reliable sources.
Suggested AI Tools	Consensus; Elicit; ResearchRabbit; Connected Papers; Perplexity
Estimated Duration	60 minutes

B. Content for STEM HE Teachers

Highlights

- AI-generated information should not be accepted without verification, as generative AI predicts likely text rather than retrieving verified facts.
- Fact-checking is a core AI literacy skill, combining critical thinking, information literacy, source evaluation, and evidence-based reasoning before AI outputs are used in education.
- AI can generate several types of factual errors, including hallucinated facts, fabricated citations, incorrect statistics, false quotations, outdated information, and statements that lack important context.
- The level of verification depends on the risk of the information. High-risk domains, such as healthcare, engineering, environmental data, public policy, and safety, require particularly detailed fact-checking.
- Effective fact-checking requires verifying claims using independent and authoritative sources, checking that references are genuine, information is current, and important context has not been omitted.
- Established information literacy frameworks, such as the SIFT Method and the CRAAP Test, provide practical approaches for evaluating the credibility and reliability of AI-generated information before it is used or shared.

Introduction

The increasing integration of AI into education has fundamentally changed the way teachers prepare lessons, develop instructional materials, assess student work, and support learning. Generative AI systems such as ChatGPT, Claude, Gemini, Microsoft Copilot, and similar large language models (LLMs) are capable of producing coherent explanations, summaries, lesson plans,

quizzes, programming code, and even scientific literature reviews within seconds. Their ability to generate human-like text has made them valuable educational assistants, reducing the time required for routine tasks and enabling teachers to focus on pedagogical activities.

However, despite their impressive capabilities, generative AI systems are not designed as factual databases. Their primary objective is to generate statistically probable sequences of words rather than retrieve verified facts from authoritative sources. As a result, AI-generated responses may contain inaccuracies, fabricated references, outdated information, or misleading interpretations presented with a high degree of confidence. These limitations create significant challenges, particularly in educational environments where teachers are responsible for providing accurate, evidence-based information to students.

For this reason, fact-checking AI-generated information has become one of the core competencies of AI literacy [1]. Rather than accepting AI outputs at face value, educators must develop the ability to critically evaluate the accuracy, reliability, completeness, and credibility of AI-generated content before incorporating it into teaching materials or recommending it to learners.

Fact-checking should therefore be viewed not merely as a technical skill but as a component of digital information literacy, critical thinking, and responsible AI use. Educators who understand both the strengths and limitations of generative AI are better positioned to use these technologies safely while simultaneously teaching students how to become critical consumers of AI-generated information.

Why AI-Generated Information Cannot Be Automatically Trusted

Unlike traditional search engines that retrieve existing documents, generative AI systems operate by predicting the most likely sequence of words based on patterns learned during their training process [2]. During training, large language models analyze enormous collections of books, websites, articles, and other textual resources. Through this process they learn statistical relationships between words, phrases, and concepts rather than storing verified facts.

Consequently, an AI model does not "know" whether a statement is true or false. Instead, it estimates which sequence of words is most likely to follow a user's prompt. While this mechanism often produces remarkably accurate responses, it also explains why AI occasionally generates convincing but entirely incorrect information.

Several factors contribute to factual inaccuracies [3]:

- incomplete or outdated training data;
- ambiguous or poorly formulated prompts;
- conflicting information within the training corpus;
- lack of access to current events or newly published research (for models without real-time retrieval);
- probabilistic text generation rather than factual verification.

One of the best-known consequences of this process is the phenomenon known as an AI hallucination [4]. Hallucinations occur when an AI system confidently generates information that appears plausible but has no basis in reality. These may include invented scientific papers, fabricated quotations, imaginary organizations, incorrect statistics, or non-existent legal regulations. For example, an AI system may produce a citation that follows the correct academic formatting style and includes realistic author names, journal titles, publication years, and page numbers, despite the referenced article never having existed. Such fabricated references are particularly problematic in education because they may easily pass unnoticed unless carefully verified.

It should be noted, that hallucinations are not software bugs but an inherent consequence of the probabilistic nature of the LLMs. Therefore, educators should approach AI-generated information as draft material requiring verification rather than authoritative knowledge.

Understanding Fact-checking

Fact-checking is the systematic process of verifying whether claims correspond to reliable evidence obtained from trustworthy and independent sources. Within the context of AI-generated information, fact-checking extends beyond simply asking whether an answer is correct. It also includes evaluating the quality of supporting evidence, identifying missing context, assessing the credibility of cited sources, and determining whether the information is sufficiently current for the intended purpose.

Effective fact-checking addresses several key questions:

- Is the factual claim accurate?
- Can the statement be confirmed by independent sources?
- Are the cited references genuine?
- Is the information up to date?
- Has important context been omitted?
- Does the response distinguish facts from opinions or predictions?
- Is uncertainty communicated appropriately?

Effective FACT-CHECKING

addresses several key questions:



Fact-checking therefore combines elements of critical thinking, source evaluation, information literacy, and evidence-based reasoning [5]. For educators, it becomes an essential quality assurance process before AI-generated materials are incorporated into classroom activities, presentations, assignments, or assessments.

Types of AI-Generated Information That Require Verification

Not all AI-generated content carries the same level of risk. The amount of verification required depends largely on the consequences of incorrect information.

Low-risk information

Some responses involve general knowledge that is relatively stable over time. Examples include definitions of technical terms, grammar explanations, mathematical formulas, programming syntax, or descriptions of well-established scientific concepts. Although occasional inaccuracies may still occur, these topics are generally easier to verify.

Medium-risk information

Educational summaries, comparisons of technologies, historical explanations, literature reviews, or recommendations for teaching methods require greater scrutiny. AI may oversimplify complex topics, omit important details, or present subjective interpretations as objective facts.

High-risk information

Certain categories of information should always undergo rigorous verification because inaccuracies may have significant consequences. These include medical advice, legal interpretations, financial recommendations, engineering calculations, public policy information, scientific findings, environmental data, and safety procedures. In these cases, AI should never be considered the final authority.



Types of AI-Generated Information That Require Verification



Not all AI-generated content carries the same level of risk. The amount of verification required depends largely on the **consequences of incorrect information**.



Low-risk information

Some responses involve general knowledge that is **relatively stable** over time. Examples include definitions of technical terms, grammar explanations, mathematical formulas, programming syntax, or descriptions of well-established scientific concepts. Although occasional inaccuracies may still occur, these topics are **generally easier to verify**.

Examples:

- ✓ Definitions (e.g., "photosynthesis")
- ✓ Grammar explanations
- ✓ Mathematical formulas
- ✓ Programming syntax
- ✓ Well-established scientific concepts



Medium-risk information

Educational summaries, comparisons of technologies, historical explanations, literature reviews, or recommendations for teaching methods require **greater scrutiny**. AI may oversimplify complex topics, omit important details, or present subjective interpretations as objective facts.

Examples:

- ✓ Educational summaries
- ✓ Comparisons of technologies
- ✓ Historical explanations
- ✓ Literature reviews
- ✓ Teaching method recommendations



High-risk information

Certain categories of information should always undergo rigorous verification because inaccuracies may have **significant consequences**. These include medical advice, legal interpretations, financial recommendations, engineering calculations, public policy information, scientific findings, environmental data, and safety procedures. In these cases, AI should **never** be considered the final authority.

Examples:

- ✓ Medical advice
- ✓ Legal interpretations
- ✓ Financial recommendations
- ✓ Engineering calculations
- ✓ Public policy information
- ✓ Scientific findings
- ✓ Environmental data
- ✓ Safety procedures



The higher the potential impact of incorrect information, the more careful the verification process should be.



Co-funded by
the European Union

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the National Agency of Erasmus+ Programme in Lithuania – Education Exchanges Support Foundation (SMPF). Neither the European Union nor the SMPF can be held responsible for them.



Common Types of AI Errors

Understanding typical AI errors enables educators to recognize situations where additional verification is necessary [6].

Hallucinated facts

The AI generates information that sounds convincing but is entirely fictional.

Example:

"The International Association of Digital Education established AI Literacy Standards in 2019."

Although plausible, such an organization or publication may not exist.

Fabricated citations

One of the most common problems in academic writing is the generation of non-existent references.

AI may produce a citation such as:

Smith, J., & Brown, L. (2023). Artificial Intelligence in Modern Education. Journal of Educational Innovation.

Despite appearing authentic, neither the journal nor the article may exist.

Incorrect numerical information

Statistics are frequently generated without proper evidence. For example:

"Seventy-eight percent of teachers regularly use AI in classrooms."

Unless accompanied by a verifiable source, such statistics should not be trusted.

False quotations

Large language models often attribute statements to famous individuals that they never actually said. For example:

"Albert Einstein once said, 'Creativity begins where algorithms end.'"

Searching reliable quotation databases reveals that no such quotation exists.

Outdated information

Rapidly changing topics such as AI technologies, legislation, software versions, or public health recommendations may no longer reflect current knowledge.

Missing context

A statement may be technically correct while still misleading. For example:

"Solar energy is the cheapest source of electricity."

This statement depends on geographical location, climate conditions, storage requirements, infrastructure costs, and the methodology used in economic comparisons.

Without this context, learners may develop an oversimplified understanding.



Co-funded by
the European Union

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the National Agency of Erasmus+ Programme in Lithuania – Education Exchanges Support Foundation (SMPF). Neither the European Union nor the SMPF can be held responsible for them.



Common Types of AI Errors

Understanding typical AI errors enables educators to recognize situations where additional verification is necessary

Hallucinated facts

The AI generates information that sounds convincing but is entirely fictional.

Example:
"The International Association of Digital Education established AI Literacy Standards in 2019."

Although plausible, such an organization or publication may not exist.

Fabricated citations

One of the most common problems in academic writing is the generation of non-existent references. AI may produce a citation such as:

Example:
"Smith, J., & Brown, L. (2023). Artificial Intelligence in Modern Education. Journal of Educational Innovation."

Despite appearing authentic, neither the journal nor the article may exist.

Incorrect numerical information

Statistics are frequently generated without proper evidence. For example:

Example:
"Seventy-eight percent of teachers regularly use AI in classrooms."

Unless accompanied by a verifiable source, such statistics should not be trusted.

False quotations

Large language models often attribute statements to famous individuals that they never actually said. For example:

Example:
"Albert Einstein once said, 'Creativity begins where algorithms end.'"

Searching reliable quotation databases reveals that no such quotation exists.

Outdated information

Rapidly changing topics such as AI technologies, legislation, software versions, or public health recommendations may no longer reflect current knowledge.

Missing context

A statement may be technically correct while still misleading. For example:

Example:
"Solar energy is the cheapest source of electricity."

This statement depends on geographical location, climate conditions, storage requirements, infrastructure costs, and the methodology used in economic comparisons.

Without this context, learners may develop an oversimplified understanding.

Frameworks for Evaluating AI-Generated Information

Several well-established frameworks from information literacy can be applied directly to AI-generated content.

The SIFT Method

Developed by digital literacy expert Mike Caulfield, the SIFT method encourages critical evaluation before accepting information [7].

- Stop - Pause before believing or sharing AI-generated content.
- Investigate the source - Identify where the information originates and whether the source is trustworthy.
- Find better coverage - Search for independent confirmation from multiple reliable sources.
- Trace claims to the original source - Whenever possible, consult the original publication rather than relying on summaries or secondary interpretations.

The SIFT method is particularly valuable because it emphasizes verification before engagement, reducing the spread of misinformation.

The CRAAP Test

The CRAAP framework evaluates information according to five criteria [8]:

- Currency – Is the information current?
- Relevance – Does it address the intended purpose?
- Authority – Who produced the information?
- Accuracy – Is it supported by evidence?
- Purpose – Why was the information created?

Originally developed for evaluating online information, this framework is equally applicable to AI-generated outputs.



Co-funded by
the European Union

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the National Agency of Erasmus+ Programme in Lithuania – Education Exchanges Support Foundation (ŠMPF). Neither the European Union nor the ŠMPF can be held responsible for them.





Frameworks for Evaluating AI-Generated Information

Several well-established frameworks from information literacy can be applied directly to AI-generated content.



The SIFT Method

Developed by digital literacy expert Mike Caulfield, the SIFT method encourages critical evaluation before accepting information.

S

II

Stop

- Pause before believing or sharing AI-generated content.

I

🔍

Investigate the source

- Identify where the information originates and whether the source is trustworthy.

F

📄

Find better coverage

- Search for independent confirmation from multiple reliable sources.

T

🔍

Trace claims to the original source

- Whenever possible, consult the original publication rather than relying on summaries or secondary interpretations.



The SIFT method is particularly valuable because it emphasizes verification before engagement, reducing the spread of misinformation.



The CRAAP Test

The CRAAP framework evaluates information according to five criteria.

C

📅

Currency

- Is the information current?

R

🎯

Relevance

- Does it address the intended purpose?

A

👤

Authority

- Who produced the information?

A

🛡️

Accuracy

- Is it supported by evidence?

P

🎯

Purpose

- Why was the information created?



Originally developed for evaluating online information, this framework is equally applicable to AI-generated outputs.

References

- [1] Zhang, Y., Fang, W., & Pan, W. (2026). Fact-Checking Misinformation in the Age of AI: Audience Responses to Human, AI, and Hybrid Sources. *International Journal of Human-Computer Interaction*, 1-17.
- [2] Banh, L., & Strobel, G. (2023). Generative artificial intelligence. *Electronic markets*, 33(1), 63.
- [3] Uddin, M., Arfeen, S. U., Alanazi, F., Hussain, S., Mazhar, T., & Arafatur Rahman, M. (2026). A critical analysis of generative AI: challenges, opportunities, and future research directions. *Archives of Computational Methods in Engineering*, 33(2), 1763-1793.
- [4] Maleki, N., Padmanabhan, B., & Dutta, K. (2024, June). AI hallucinations: a misnomer worth clarifying. In *2024 IEEE conference on artificial intelligence (CAI)* (pp. 133-138). IEEE.
- [5] Wolfe, R., & Mitra, T. (2024, June). The impact and opportunities of Generative AI in fact-checking. In *Proceedings of the 2024 ACM Conference on Fairness, Accountability, and Transparency* (pp. 1531-1543).
- [6] Mahmood, A., Fung, J. W., Won, I., & Huang, C. M. (2022, April). Owning mistakes sincerely: Strategies for mitigating AI errors. In *Proceedings of the 2022 CHI conference on human factors in computing systems* (pp. 1-11).
- [7] Caulfield, M., & Wineburg, S. Learning to Verify.
- [8] Blakeslee, S. (2004). The CRAAP test. *Loex Quarterly*, 31(3), 4.

Green Decision Context

AI-generated information increasingly supports green and environmental decision-making, including:

- interpreting environmental monitoring data (air, water, soil, biodiversity)



Co-funded by
the European Union

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the National Agency of Erasmus+ Programme in Lithuania – Education Exchanges Support Foundation (SMPF). Neither the European Union nor the SMPF can be held responsible for them.



- summarizing climate change research and sustainability reports
- providing information on renewable energy technologies and circular economy practices
- explaining environmental regulations, standards, and sustainability policies
- supporting education on climate action and sustainable development

Before using AI-generated information in green contexts, users should verify:

- scientific claims against trusted environmental organizations and peer-reviewed sources
- numerical data such as carbon emissions, energy consumption, and environmental indicators
- whether environmental regulations and sustainability standards are current
- that cited reports, legislation, and scientific publications actually exist
- whether important environmental trade-offs or local conditions have been omitted

Videos/Interactive Resources

- Fact checker – Manus - <https://manus.im/playbook/fact-checker>
- Fact-checker – FullFact - <https://fullfact.ai/>



Co-funded by
the European Union

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the National Agency of Erasmus+ Programme in Lithuania – Education Exchanges Support Foundation (SMPF). Neither the European Union nor the SMPF can be held responsible for them.

