Guidelines for Elementary, Junior High and Senior High School on Using Generative Artificial Intelligence: For Teachers, Administrators, and Parents

Approved by the Ministry of Education, Republic of China, on July 1, 2024
(Taiwan Education Resource (3) Letter No. 1132702614)

In recent years, the rapid development of generative artificial intelligence (generative AI) tools and deepfake technologies has resulted in numerous opportunities and changes for society. These technologies have transformed the ways we acquire knowledge, disseminate information, and create content. However, the risks associated with their use have also increased. To assist elementary, junior high and senior high school teachers, administrators, and parents in skillfully using these technologies, enhancing their literacy in using generative AI tools, and preventing misuse and abuse, the following six guidelines are provided for reference.

1. **Recognize Potential Bias in AI-Generated Content**. The data used to train generative AI tools are based on historical records or experiences. If these training data contain biases or errors, the results produced by the generative AI tools will also be biased or incorrect. Moreover, these tools cannot independently determine the accuracy and reasonableness of their outputs. Therefore, when using generative AI tools, we must review the results to ensure their correctness and logical consistency.

2. **Understand the Reduction of Information Diversity by Generative AI Tools**. If the data used to train generative AI tools are limited by region, culture and lack of diversity, the results generated by these tools may only reflect a single cultural perspective and knowledge context, potentially reinforcing existing biases. Thus, when using generative AI tools, we must combine our own experience and critical thinking skills to evaluate the results instead of accepting the generated content uncritically.

3. **Acknowledge the Limitations of AI Content Recognition Tools.** With the rapid development of generative AI tools, there are many available that can help identify generated content. Therefore, we need to understand that these detection tools have their limitations, relying on common sense, intuition, and other evidence to determine whether content originates from generative AI tools is required.

4. **Be Aware of the Increasing Realism and Potential Misinformation of Deepfake Technology**. Deepfake technology, which uses generative AI to create fake content, can be used to manipulate images, videos, or audio materials to produce realistic videos and images, including fake news. Hence, when viewing online content, we must not readily trust unverified videos or photographs and must be aware of the possibility that this content may have been synthesized using deepfake technology. We must also assess the motives and intentions behind the creation of such content.

5. **Realize that "generative AI" Tools Might Leak Personal and Organizational Privacy and Confidential Information**. Some generative AI tools lack comprehensive legal, regulatory, and ethical oversight concerning the acquisition, storage, and use of data. Thus, personal information, sensitive messages, and confidential data may be provided when using these tools may be incorporated into training databases and used in future responses to others. When using generative AI tools, we must strictly and carefully assess the provided information for any content related to confidentiality, privacy, or sensitivity. In this way, we protect the privacy and confidentiality of individuals and organizations.

6. **Avoid Overreliance on Generative AI Tools to Prevent Intellectual Property Infringement and Violation of Academic Ethics**. When we use generative AI tools to generate teaching materials, test questions, lesson plans, and other educational content, we must scrutinize the content and terminology to ensure they meet educational standards and usage norms. We must regulate the timing and manner of using generative AI in academic contexts and remind students that improper use could infringe on others' intellectual property rights and violate academic ethics. Failure to provide proper citations leads to the plagiarism of others' ideas.

Though development of generative AI has resulted in substantial convenience for our lives and has been widely applied in various contexts, it carries risks and challenges. In this digital age, we must maintain a high level of vigilance regarding information sources, avoid easily-believed unverified information, and learn how to identify false information; improve our critical thinking abilities, critically analyzing and evaluating the content generated by generative AI tools to avoid being misled; and adhere to ethical and legal standards, ensuring that the use of generative AI tools does not violate social norms and information ethics. Finally, we must strengthen our digital literacy. Only by strengthening digital literacy can we enjoy the considerable convenience of technological advancements, minimize the risks, and reduce the negative effects of generative AI.

Appendix 1

Guidelines for Elementary, Junior High and Senior High School on Using Generative Artificial Intelligence: For Teachers, Administrators, and Parents Version

PARADIGMS

1. **Recognize Potential Bias in AI-generated Content**. When we ask a generative AI tool to suggest a travel itinerary, if the system's database lacks the correct information about local climate, geographical locations, social customs, and cultural restrictions, the provided content may be a mere synthesis of various online travel blogs. This can result in an itinerary which is out of your way, out-of-season, or even includes nonexistent attractions.

2. **Understand the Reduction in Information Diversity by Generative AI Tools**. When we consult a generative AI tool for legal or cultural inquiries, the answers may be based on the laws and cultural practices of the developer's country. Similarly, if we request a generative AI tool to generate an image of a bride, it may produce an image of a Western woman in a white gown, rather than reflecting the diverse cultural customs of the user's locale, such as varied skin tones or wedding attire.

3. **Acknowledge the Limitations of AI Content Recognition Tools**. When using AI content recognition tools, we can quickly compare the similarity between two or more articles, but these comparisons should only be employed as preliminary references. Determining whether an article constitutes textual or conceptual plagiarism, or whether it was entirely by generated AI tool, requires a manual comparison of the sources, close reading, and comprehension of the contents.

4. **Be Aware of the Increasing Realism and Potential Misinformation of Deepfake Technology**. Numerous online videos involve well-known individuals giving speeches or encouraging investments. When encountering such content, we must carefully verify the information and its sources. In the rapidly evolving online environment, deepfake technology may integrate the faces and voices of these individuals into entirely false and defamatory videos without their consent or knowledge.

5. **Realize that "generative AI" Tools Might Leak Personal and Organizational Privacy and Confidential Information**. When we ask generative AI tools questions by using personal or company documents or trade secrets without fully understanding the principles and regulations that govern the operations of these AI tools, then we risk incorporating the documents or Code Confidentiality into the training databases of these AI tools. When other users subsequently ask similar questions, then the generative AI tool can respond with information from the stored database, which could lead to breaches of confidential, personal or academic information, and therefore security vulnerabilities.

6. **Avoid Overreliance on Generative AI Tools to Prevent Intellectual Property Infringement and Violation of Academic Ethics**. When we use generative AI tools to generate project plans, if some terms or phrases are not idiomatic, they must be corrected. Similarly, when these generative AI tools are used to create exam questions, the appropriateness of these questions and the accuracy of the answers must be reviewed.

Appendix 2

List of AI-generated Content Detection Tools

Name of tool	Features	Supported content types	Source
Content at Scale	The development team at Content at Scale believes that texts produced by generative AI tools such as ChatGPT, Claude, and Gemini leave certain distinctive traces in terms of word usage or syntactic structures. The development team has analyzed a vast number of blog articles, Wikipedia entries, academic papers, and texts generated by multiple large language models to enhance learning and training processes. This approach helps determine the likelihood that a piece of text has been generated by generative AI tools.	Text	https://contentatscale.ai/ai-content-detector/
CopyLeaks	The developers' official statement claims a detection accuracy rate of over 99% and a false positive rate of 0.2%. CopyLeaks can detect AI-generated content in over 30 languages in texts generated by ChatGPT, Gemini, and Claude. Additionally, CopyLeaks supports browser extensions and integration into websites or learning management systems.	Text	https://copyleaks.com
GPT Detector	The primary function of the GPT Detector is to evaluate the proportion of content that is likely generated by GPT-3, GPT-4, or ChatGPT. This tool is available for free use by general users; however, to prevent misuse, a daily usage limit is set. Additionally, the developers explicitly state that the data evaluated are encrypted and not stored.	Text	https://x.writefull.com/gpt-detector
GPTZero	GPTZero analyzes AI-generated text at the document, paragraph, and sentence levels, providing the probability that the content is AI-generated and confidence interval for the results. The tool can also identify AI-generated content across a series of text files, comparing up to 50 files simultaneously with a total capacity not exceeding 15MB. Each text file can contain up to 50,000 characters.	Text	https://gptzero.stoplight.io
Hive Moderation	Hive Moderation can detect AI-generated text, images, videos, and audio, and can be integrated as a Chrome extension. The detection results are presented as percentages, and the tool highlights the areas which are likely to contain AI-generated content based on the detection outcomes.	Text, images & videos	https://hivemoderation.com/ai-generated-content-detection
Hugging Face	Hugging Face is a developer of AI and machine learning tools. In 2019, before the advent of ChatGPT, they had established an AI content detection website. Users must input approximately 50 words for the tool to provide the probability that the content was generated by artificial intelligence.	Text	https://huggingface.co/learn/nlp-course/zh-TW/chapter4/2

Name of tool	Features	Supported content types	Source
Scribbr	Scribbr is suitable for detecting content generated by ChatGPT, GPT-4, and Gemini. The tool verifies the originality and authenticity of text by examining features such as sentence structure or length, word choice, and predictability. Scribbr has a user-friendly interface and can perform unlimited text detections for free without creating an account. For accuracy, the provided text content must be between 25 to 500 words.	Text	https://www.scribbr.com/ai-detector/
Turnitin	Turnitin is the most widely used online plagiarism detection system globally. Turnitin allows users to upload personal dissertation files and automatically calculates the percentage of text that is similar to other sources within minutes. The system highlights similar sections and identifies potential original sources, enabling researchers to independently verify the originality of their work. This process enhances the credibility of dissertations.	Text	https://www.turnitin.tw/
Winston [ai]	This tool can detect content generated by ChatGPT, GPT-4, Gemini, and other generative tools, with an officially stated accuracy of 99.98%. Currently, Winston [ai] supports multiple languages, such as English, French, German, and Spanish. The tool also supports Optical Character Recognition technology, allowing it to extract text from images or scanned documents, including handwritten content.	Text & images	https://gowinston.ai
Writer	Writer can evaluate text to determine if it was generated by AI, with a minimum word count of 60 words and a maximum of 5,000 words per examination. If a person writes using similar word sequences to those in Writer's database, the tool can detect these similarities. Although the tool is not 100% accurate, it can indicate the likelihood that content was generated by AI.	Text	https://writer.com/ai-content-detector/
PPVS	PPVS analyzes linguistic features within a text to determine if an article was produced by generative AI. PPVS can identify texts produced by AI language tools such as GPT-4, GPT-3, GPT-2, BERT, and Jasper. Additionally, the tool can detect whether translation software was used in the creation of an article.	Text	https://ai.ppvs.org/

1. Notes: Tools are listed in alphabetical order.
2. With the exception of Hugging Face, all tools require a subscription fee (although free versions are available).
3. More generative AI content detection tools are likely to be developed. This table only provides common tools for reference.