WORLD HEALTH ORGANIZATION

Information Note: 09/2025

Subject: Guidance on using advanced Artificial Intelligence (AI) in the workplace

Distribution: All WHO workforce and contractors working under WHO agreements

Date: 27 February 2025

Comments: This Information Note replaces Information Note 20/2023 Navigating the issues and risks of generative artificial intelligence (AI) in the workplace

PURPOSE

The purpose of this Information Note is to provide guidance to members of the workforce on the **responsible and safer use of advanced AI, such as open-source or proprietary General Purpose Artificial Intelligence (GPAI), and Generative Artificial Intelligence (GenAI) tools, in the workplace**. This guidance is also relevant to work performed by contractors under WHO agreements.

WHO is committed to creating an empowering and accountable work environment where members of the workforce are enabled to use advanced AI technology, to enhance their productivity, efficiency and achieve their best work. This guidance outlines the expectations for using AI responsibly and effectively, ensuring alignment with WHO core values and code of ethics. The primary focus is on the use of AI applications associated with GPAI and GenAI on WHO devices and/or personal devices when used for work purposes.

BACKGROUND

AI is an umbrella term for technologies that allow machines to perform functions that normally require intelligence, such as learning, reasoning, and problem-solving. Within this field, General Purpose AI (GPAI) and Generative AI (GenAI) have emerged as two prominent interconnected branches that are rapidly evolving and transforming modern work. The technology is trained using vast datasets, such as text, images, audio, and video content that is publicly available and sourced from the internet.

GPAI focuses on performing a wide range of tasks, such as automation, recommendations, analysis across domains, GenAI specializes in generating new content — such as text, audio, visual imagery, or video—based on a given prompt or input data. Many AI tools combine both, with GPAI handling diverse functions and GenAI powering creative tasks like text, image, or code generation. Many AI applications emerging today combine both GPAI and GenAI, such as digital assistants and productivity tools.

Key examples include:

- **Digital assistants and productivity:** such as ChatGPT, Copilot, Copilot M365, Google Gemini, Claude.ai, LLaMA, Amazon CodeWhisperer, Otter.ai, Fireflies.ai (meeting notes, summaries).
- **Text, speech, video, and code generation:** such as Jasper AI, Copy.ai, Writesonic, Synthesia, Pictory, Descript, GitHub Copilot.
- **Image and design generation:** such as DALL·E, Midjourney, Stable Diffusion, Adobe Firefly, Runway ML, Canva AI, Fotor AI.
- **Music and audio generation:** such as AIVA, Soundraw, Amper Music, Murf.ai, ElevenLabs.
- **Translation and interpretation:** such as Microsoft Translator, DeepL, Google Translate, Wordly, Unbabel.

The sharing of content with public AI tools can potentially make confidential or sensitive information available to the public as the tool may train its model on the data shared. By following this guidance, the WHO workforce will be able to better understand how to use AI responsibly in the workplace while maximizing its benefits.

IMPORTANT GUIDANCE

Use case examples

To illustrate practical applications of the guidelines, a list of **example of AI do’s and don'ts** in the workplace has been created. This list is not exhaustive and may evolve over time.

The use cases are divided into two categories:

- **Internal-facing:** AI used for internal operations, communications, and processes.
- **External-facing:** AI used when interacting with Member States, Donors, Partners, Civil Society, vendors, media or other external partners.

Some scenarios may overlap or change between internal and external scenarios, so stay alert to these instances and ensure you adapt.

1. Ensure contractors adhere to WHO's AI guidelines for AI-supported work.

- When overseeing work by contractors under WHO agreements, responsible officers should discuss WHO’s AI guidelines with them and ensure they comply with WHO’s standards, maintaining the integrity of AI applications.
- Any planned use of AI in the development of the work should be discussed at the outset and spelt out in the Terms of Reference. This includes the platform used, license agreements in place, steps to mitigate copyright and ethical risks, development and retention of the prompting instructions.
- Contractors should document how AI has been used in the development of the work produced when the work is delivered.
- Contractors should have a professional license in place with the AI platform used. This provides them with (i) the right to use the platform for commercial purposes (ii) the ability to develop and use a closed AI model, (iii) develop detailed queries non-publicly.
- Contractors should ensure that no WHO materials are released into the public AI models.

2. Use AI tools covered by WHO Enterprise Agreements

- Use GPAI and GenAI tools (e.g., Copilot, Copilot M365) that are secured under WHO Enterprise Agreements and comply with policies, data exchange agreements, and best practices for data security to minimize risks associated with data breaches, unauthorized access, and misuse.
- Ensure tools are backed by Enterprise Data Protection (EDP) commitments and meet WHO standards for security, compliance, and privacy for handling sensitive and confidential WHO content. Be aware that certain AI plugins linked to public platforms are available for browser use. These plugins, when enabled, could inadvertently expose confidential or sensitive information to public large language models (LLMs) or other external platforms.

- Public GPAI and GenAI tools not covered by WHO Enterprise Agreements can be used for public content, as appropriate.

3. Safeguard WHO data

- Ensure compliance with WHO's **privacy policy** and **data protection policy** and **data principles**.
- Do not share identifiable information such as photos, personal, sensitive, confidential data or proprietary information belonging to third parties.
- Avoid sharing any information not yet approved for public access into open-source or proprietary AI tools (this includes integrating unapproved AI tools into meetings for tasks such as note-taking or summarizing meeting minutes).
- Be aware of how data is being collected, used, and who will have access to it.
- Be cautious when agreeing to external terms and conditions to prevent unintentional transfer of ownership.

4. Commit to ethical AI

- Use AI tools responsibly and exclusively for legitimate and ethical purposes. Do not generate any content that is malicious, inappropriate, or illegal, such as malware, deepfakes, harmful or toxic material, or any other content unsuitable for a professional work environment.
- Follow the **WHO Code of Ethics and Professional Conduct** when using AI-powered solutions. Engage in respectful, ethical, and appropriate practices, such as obtaining consent from impacted participants and being mindful of where data is stored by third-party vendors.

5. Ensure transparency in AI use

- WHO-released or endorsed apps **must clearly communicate the use of AI systems**. This transparency helps users understand when and how AI is being utilized, fostering trust and accountability.
- AI use in WHO-released or endorsed apps must undergo appropriate reviews to assess quality, safety, cybersecurity, and other privacy and security best practices.

Potential shortcomings including the risk of **hallucination** or erroneous output should be disclosed, as well as methods undertaken to mitigate and minimize risks.

6. Monitor for potentially harmful AI-generated material (Validation PROCESS)

- Be aware that AI solutions may contain inherent biases, harmful, stereotypical, discriminatory, or otherwise offensive language or images. Understand that these biases can be influenced by the data used for training the algorithm or due to deliberate efforts and can propagate these biases and language in their outputs.

- AI should assist, not replace, human judgment. Human oversight and judgment are essential to validate AI recommendations and intervene when necessary to correct or mitigate any issues arising from AI outputs.

7. Cross-check AI-generated output (VALIDATION AND CROSS CHECKING / AVOIDING HALLUCINATIN BY HUMAN REVIEW, FACT CHECKING)

- Always review AI-generated content thoroughly. Verify sources for validity, accuracy, completeness, and safety.

- Understand that AI solutions can sometimes produce convincing but inaccurate recommendations or misinformation (often referred to as hallucinations).

8. Ensure appropriate content and media attribution

- When citing sources, always follow standard scientific practices by attributing only primary sources. Avoid using AI-generated content as references. Ensure your content is accurate, respects copyright laws and maintain integrity by avoiding plagiarism.

- In the methods section of a published document, include a statement disclosing the use of AI tools for content research. Verify primary sources before publication, and if manual verification isn't feasible, use plagiarism detection software to ensure the content's originality and reliability.

- ~~AI should not be used to create realistic images or videos of fictional characters, places or situations. Photos and videos published by WHO must show real people in genuine situations to maintain trust and integrity. Where AI is used to generate graphic elements or non-realistic images, this should be clearly labelled.~~

- ~~Obtain consent for identifiable individuals, except for crowds in public spaces in visual or voice context. Avoid discriminatory scenarios, negative stereotypes or~~

~~stigma scenario without disclaimers and uphold gender equality. Additionally, any media containing identifiable and recognizable people and other materials should not be uploaded to publicly available sites (LLMs), where datasets can be used to train or improve public AI models.~~

9. ~~Be cybersecurity conscious~~

- ~~When using AI tools, it is important to follow WHO cybersecurity guidelines to protect WHO digital assets (refer to cybersecurity resource portal). AI-based systems, like many information systems, can be exploited by cybercriminals to create advanced malware and conduct social engineering attacks, like phishing and spear-phishing, to steal confidential information or spread harmful links. Ensure you are not redirected to a spoof site mimicking a genuine AI interface.~~

10. Follow secure file sharing practices

- As we continue to use various tools and technologies in our daily work, it's important to stay attentive to our file-sharing practices. While this is not a new risk, the integration of AI makes it easier for sensitive information to be accessed, processed, or misused if files are not handled carefully. AI's capabilities can quickly analyze and disseminate information, amplifying the impact of any accidental or unauthorized sharing.
- To mitigate these risks, only share files with those who need access, verify permissions (e.g. SharePoint permissions), beforehand, and be cautious when sharing sensitive or confidential information. Permissions to drafts and old versions should be restricted to avoid the risk of misuse of its contents.

11. Maintain a clear distinction between your work and personal accounts and data (where possible)

- When registering for public internet services or creating online accounts (e.g. for AI tools and services, as well as non-AI related services) that are non-work related, it is recommended to avoid using your work email or credentials (or vice versa). This precaution helps protect against data breaches and prevents unauthorized access to official systems through your connection.

CONCLUDING REMARKS

Following these guidelines will help WHO leverage AI to enhance efficiency and productivity in our workplace, while maintaining a base level of safety and security. This guidance will be regularly reviewed and updated to keep up with the rapidly evolving landscape of AI and advanced technologies. The WHO workforce is encouraged to keep their AI skills and knowledge current by staying informed about advancements, sharing best practices, experimenting responsibly, and engaging in continuous learning activities.