AI FORUM

Year: 2026 | Edition: 6 |
Cambodia | 11 February 2026 |

INSIGHTS

Generative AI Use: A Task-Model-Mode-Verification Framework

LIM Sengkhoun [1], MD, MHPE, International FAIMER Fellow

[1] ***LIM Sengkhoun*** *is a Health Advisor at GIZ Cambodia, Adjunct Assistant Professor at University of Puthisastra, Part-time Educator at University of Health Sciences, and Research Fellow at the AI Forum.*

Asking good questions is a starting point for working with generative AI (GenAI), but it is not enough. Even well-crafted prompts with context can produce poor outcomes if wrong models are chosen, tasks are inappropriately approached, or outputs are accepted without adequate verification. Many common GenAI failures - confident errors, plausible misinformation, and uncritical reuse of outputs - arise not from insufficient prompting (Milam & Gulli, 2025), but from misalignment between the tasks, models, working modes, and levels of human oversight applied.

GenAI use therefore requires moving beyond a single-tool mindset to a deliberate system of use. Different tasks need different model capabilities, different operational modes prioritize speed or analytical depth, and every output requires verification proportional to its risks.

This paper introduces a Task-Model-Mode-Verification framework to support responsible and effective GenAI use by aligning task purpose with model selection, thinking modes, and verification practices within a coherent decision-making process.

Choosing Right Models: Fit for Purpose Matters

Not all GenAI models are designed to do the same kind of work. Some are optimized for general reasoning, explanation, and drafting, while others specialize in retrieving up-to-date information, linking to external sources, or analyzing user-provided documents. Model selection should therefore be primarily guided by task purpose, not convenience, familiarity, or popularity.

Different categories of models tend to demonstrate distinct strengths. Conversational reasoning models support conceptual explanation, drafting, and reasoning through ideas. Search-integrated systems are more suitable when tasks require locating recent information or identifying traceable sources. Some models are optimized for synthesizing extensive user-provided materials across long documents or multiple data sources, while others function effectively within productivity ecosystems or multilingual workflows. These differences reflect architectural characteristics - such as context-window capacity, retrieval integration, and document-processing - rather than fixed product distinctions.

Problems arise when users expect a single model to perform all tasks equally well (Su et al., 2025). For example, relying on conversational models may generate fluent explanations while

lacking robust evidence retrieval, whereas search-focused systems may identify sources without deeper conceptual reasoning. Effective GenAI practice begins with a fundamental question: What type of work am I trying to accomplish - reasoning, retrieval, synthesis, drafting, or verification? Skilled users treat tool selection as dynamic and may switch models within a single workflow. As GenAI evolves rapidly, model strengths reflect shifting tendencies rather than fixed capabilities, making task-driven and adaptive model selection essential.

Working Modes: One Model, Different Ways of Thinking

Even within a single model, performance varies greatly depending on how it is used. Modern GenAI systems operate through different working modes - some optimized for response speed and fluency, others for depth, analysis, and synthesis. These modes may be explicitly labeled in user interfaces or implicitly activated through prompt design, interaction styles, and task sequencing.

This distinction parallels human cognition. Professional decision-making often involves shifting between fast, intuitive thinking and slower, analytical reasoning depending on context. Rapid judgment is effective in familiar situations, while complex or high-stakes decisions require careful analysis. This contrast is widely described as fast (System-1) and slow (System-2) thinking (Kahneman, 2011).

Similarly, rapid GenAI working modes produce immediate, fluent responses useful for brainstorming, early drafting, summarization, and exploratory learning. However, like fast human thinking, these modes are vulnerable to shallow reasoning, hidden assumptions, and overconfidence - particularly when problems are ambiguous or inputs are incomplete.

Analytical working modes are engaged when users request stepwise explanations, comparison of alternatives or assumption testing. These modes are better suited for research synthesis, complex learning tasks, and decision-making. Although analytical modes do not guarantee correctness, they increase transparency and enable critical evaluation.

Failures often occur when rapid modes are used for tasks requiring deeper analysis. Fluent outputs can reinforce initial impressions, reduce doubt, and discourage verification, creating misplaced confidence. Effective GenAI use therefore requires deliberate mode switching.

Rapid modes support exploration and productivity, whereas analytical modes are essential when uncertainty, risk, or accountability is high.

Verification: Human Professional Responsibility

Regardless of model choice or working mode, GenAI outputs are probabilistic and lack accountability (Rahman et al., 2026). While GenAI can critique their own outputs or generate supporting arguments, they cannot verify truth in a meaningful sense. Responsibility for accuracy therefore remains with human users, making verification a core professional competency in GenAI-augmented environments (UNESCO, 2024).

Verification should be proportionate to risk. Low-stakes tasks such as brainstorming or rough drafts require light checks for coherence and plausibility. Medium-stakes tasks, including teaching materials and professional communication, require closer examination, contextual interpretation, and source validation. High-stakes tasks (e.g., research claims, policy recommendations, clinical reasoning) need rigorous verification against original sources, professional standards and established safeguards.

Across contexts, three practical verification strategies apply. First, *check internal logic* by examining whether conclusions follow from reasoning, assumptions are explicit, and claims align with available evidence. Second, test robustness by generating counterarguments, alternative interpretations, or conditions under which conclusions might fail. Third, verify externally when consequences are significant by treating GenAI outputs as provisional leads and checking original sources for accuracy, reliability, and contextual relevance.

In professional practice, domain-specific safeguards remain essential. GenAI should support reasoning, not replace accountable judgment. Verification strengthens critical thinking and ensures that GenAI amplifies rather than substitutes human responsibility.

Conclusion

In a world of instant answers, professional judgment becomes the true differentiator. GenAI can generate ideas, arguments, and evidence at unprecedented speed, but it cannot bear responsibility for what is true, safe, or ethically right. That responsibility remains human. The professionals who will lead in a GenAI-augmented future are those who align tasks with

appropriate models, switch between rapid and analytical thinking, and verify outputs proportionate to potential consequences. Expertise is therefore defined not by access to information alone, but by the ability to use GenAI with intention, skill, and accountability - knowing when to move fast, when to slow down, and when accuracy must outweigh fluency.

Reference:

Kahneman, D. (2011). Thinking, fast and slow. *Farrar, Straus and Giroux*.

Milam, K., & Gulli, A. (2025). *Context engineering: Sessions & memory*. Google. https://www.kaggle.com/whitepaper-context-engineering-sessions-and-memory

Su, W., Long, J., Wang, C., Lin, S., Xu, J., Ye, Z., ... & Liu, Y. (2025). Towards Unification of Hallucination Detection and Fact Verification for Large Language Models. *arXiv preprint arXiv:2512.02772*.

Rahman, S. S., Islam, M. A., Alam, M. M., Zeba, M., Rahman, M. A., Chowa, S. S., ... & Azam, S. (2026). Hallucination to truth: a review of fact-checking and factuality evaluation in large language models. *Artificial Intelligence Review*.

UNESCO. (2024). *AI competency framework for teachers*. https://unesco.org/en/articles/ai-competency-framework-teachers

** Edited with ChatGPT for clarity; all views expressed are solely those of the author.*
** The pinions expressed are the author's own and do not reflect the views of AI Forum.*