

Part 3: Common Questions and Critical Reflections

This section addresses key practical and ethical considerations to help you avoid common pitfalls when working with GenAI.

Question 1: Is the “research gap” found by AI really a gap?

Common situation: AI says “a certain direction has received little research”, and you immediately take it as an original contribution.

Reflections:

- “Little research” does not automatically equate to academic value. Sometimes it is because the direction itself is unimportant, infeasible, or has already been disproven.
- A true research gap must satisfy three conditions: a. not systematically studied; b. has theoretical or practical significance; c. data are available, and you can actually do the work.
- Suggested approach: Search for the claimed gap using different keyword combinations in databases such as Google Scholar, CNKI, and Scopus. Only if all three searches find no relevant literature can you be relatively confident. Among the three conditions for a genuine gap, AI can help you identify a, but you must judge b and c yourself.

Question 2: “Does an AI tool ‘pretend’ to have read a paper?”

Yes. This is the most common hallucination problem.

When you ask AI “What does scholar X think about this issue?” it may fabricate a plausible-sounding answer, or even invent paper titles, journal names, and years.

Prevention methods:

- Ask the AI to provide a direct quotation (“Please quote the exact passage from the original paper”). If it cannot provide one, be highly suspicious.
- Copy the suspicious paper title into CNKI, Google Scholar, or other databases to verify its existence.
- Trust only the PDF content you upload yourself. Use the “upload file” function of DeepSeek, ChatGPT, or Perplexity, and let the AI answer based on the text you provide.

Question 3: Can I directly use a research title generated by GenAI?

Not recommended. AI-generated titles often:

- Tend to be mediocre and lack sharpness
- Use formulaic structures (“A study of ...”, “Based on ...”)
- May contain methods you cannot actually implement (e.g., “a neurolinguistic perspective”)

Suggested workflow:

1. Let the AI generate multiple options (descriptive, comparative, explanatory).
2. You select 2–3 that feel feasible.
3. Search those 2–3 titles in literature databases to confirm that no identical paper exists.
4. Finally, rewrite the title yourself, adding your unique data source or perspective.

Question 4: Do I need to declare the use of GenAI in my paper?

Yes, you need to declare it.

For details, please refer to the Hong Kong Polytechnic University’s guidelines on the use of generative AI: https://www.polyu.edu.hk/edc/explore-a-topic/generative-ai/?sc_lang=tc

Question 5: Does the AI's "knowledge cut-off date" affect my research?

Yes. This is a frequently overlooked limitation.

Most general-purpose models have a knowledge cut-off date, which means:

- The AI does not know papers published after that date
- The AI does not know the linguistic changes that occurred after that date
- If you are studying a very recent phenomenon, the AI may be completely unaware of it

Solutions:

1. First ask the AI: "What is your knowledge cut-off date?"
2. For literature after that date, you must search manually (Google Scholar, CNKI, etc.).
3. Upload the PDFs of new papers to the AI and let it answer based on the new information.

Question 6: How can I tell whether a GenAI tool is trustworthy?

You can evaluate a tool along these dimensions:

Evaluation dimension	Question	Ideal answer
Citation transparency	Does the tool provide direct citation links?	Yes (e.g., Consensus, Elicit)
Hallucination rate	How serious are known hallucination problems?	Low (academic-oriented tools are lower)
Data source	Which database does it rely on?	Public academic papers
Paywall	Is the free version sufficient?	Sufficient for coursework
Chinese support	How well does it support Chinese literature and queries?	CNKI AI is best; general models are secondary

Features that make a tool NOT recommended:

- Does not provide citation sources
- Claims to "write an entire paper for you"
- Unclear source or based only on a small, unknown database
- Asks you to upload large amounts of unpublished data without explaining how it will be used

Question 7: Can I directly upload collected corpora or interview transcripts for AI analysis?

Absolutely not. This is an extremely sensitive red line concerning data security and academic ethics.

Free or public versions of large AI models (e.g., the free version of ChatGPT, Doubao, etc.) typically collect user input to train future models. This means:

- If you upload unpublished original corpora or your supervisor's project data, your research results may be leaked in advance.
- If you upload interview transcripts containing real information about respondents (e.g., from a sociolinguistic survey), you will seriously violate privacy protection and academic ethics.
- If you upload complete copyrighted corpora or paid books, you may be infringing copyright.

Solutions:

1. Complete anonymisation: Before entering any text into an AI, manually delete all real names, specific locations, and other personally identifiable information.
2. Use protected platforms: Prioritise official AI tools provided by your university (e.g., PolyU GenAI). Such enterprise-level licences typically have strict privacy policies that promise “not to use user data for model training”.
3. Process sensitive data locally: For highly sensitive or unpublished core corpora, avoid using online AI tools entirely. Instead, use traditional local analysis software (e.g., NVivo, MAXQDA), or rely solely on your own in-depth analysis.

Bonus Tool: Critical Self-Check Table

After each use of GenAI to assist your research, quickly check whether your usage was appropriate using the following questions:

Question	Yes	No
Did I verify the “research gap” claimed by AI?	☐	☐
Did I verify all references provided by AI?	☐	☐
Did I personally read the abstract or full text of at least three core papers?	☐	☐
Did I perform a feasibility check on the AI-generated title?	☐	☐
Do I know the knowledge cutoff date of the tool I used?	☐	☐
Am I prepared to declare my use of GenAI in my paper?	☐	☐

Conclusion

GenAI is a powerful assistant, but its outputs require your critical judgment to validate. The seven questions and the self-check table in Part 3 of this guide are meant to help you stay alert while using these tools:

AI can help you find directions, but the final judgment, choices, and responsibility are all yours.



Video Summary 3: Common Questions and Critical Reflections

Full text Ended