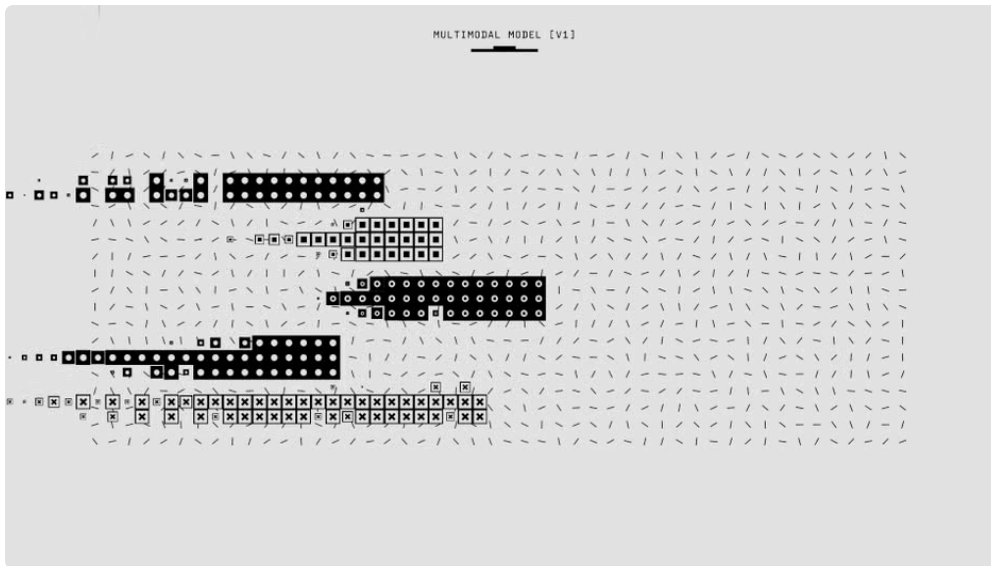


As a technology risk lead working across ultra-high net worth (UHNW) family offices in Canada, I'm constantly evaluating how AI tools like ChatGPT and Google's latest Gemini model integrate into investment research workflows. The allure of rapid summaries, parsing dense reports, and generating preliminary due diligence is undeniable. Yet, behind the convenience lies a fundamental challenge: **AI hallucinations** — persuasive but potentially inaccurate outputs — that can mislead decision-makers without proper attribution or traceability.



This article explores practical steps and considerations to **request citations**, extract **source links**, and obtain verifiable **quotes** when using prominent AI tools. We especially focus on family office use cases, where private market data scarcity compounds risks, and why forcing transparent sourcing remains crucial for effective governance.

Understanding AI Hallucinations: The Risk You Can't Ignore

AI hallucinations refer to instances where generative models produce outputs that appear factual, authoritative, and confident — yet are incorrect, fabricated, or misleading. Unlike simple errors, hallucinations can be highly polished prose or plausible-sounding data points that are entirely artificial. This subtlety makes them especially dangerous in domains like investment research, where trust and accuracy are paramount.

Why Are AI Hallucinations a Concern in Family Offices?

- **Preliminary Research:** Family office analysts often use AI tools to quickly gather background info on new asset classes, sectors, or managers. An incorrect fact dropped without a source can propagate unchecked.
- **Parsing Complex Reports:** AI helps summarize long documents, but without citations, it's impossible to validate whether key claims were accurately represented or invented.
- **Private Markets Scarcity Amplifies Risk:** Public market data is broadly available for verification; private equity or venture capital data is thin and fragmented, making hallucinations harder to detect.

Faithfulness vs. Factuality: What the AI Is Really Giving You

A critical distinction often lost in hype is between an AI output's *faithfulness* versus *factual accuracy*. Faithfulness means the output accurately reflects the input text or the data it was trained on. Factual accuracy means the information corresponds to real-world truth.

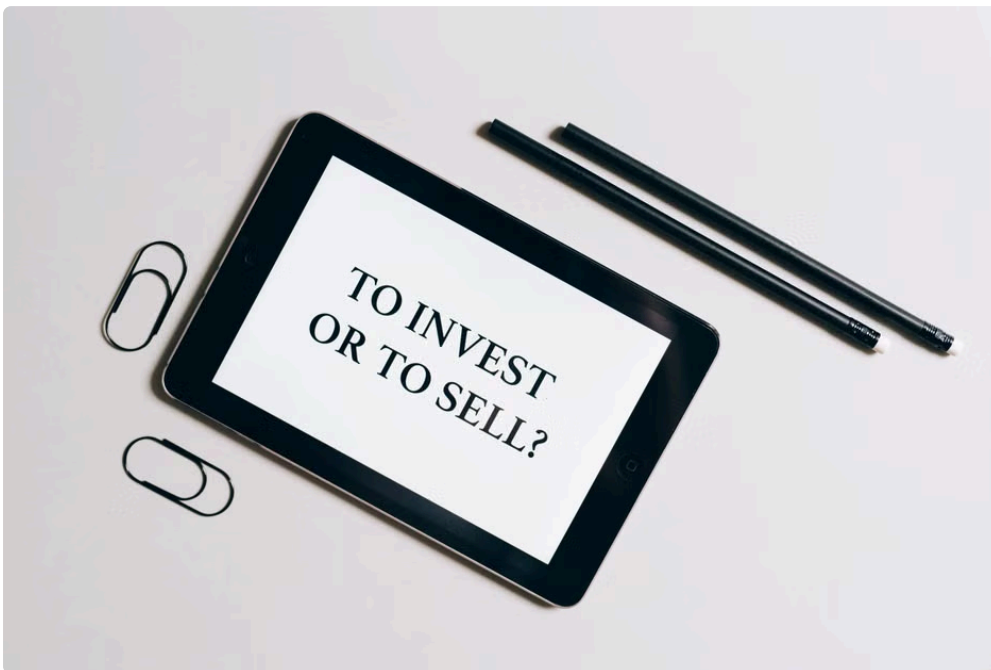
For example, an AI summarizing a manager's strategy based on their pitch deck might faithfully restate points, even if the deck contains optimistic or misleading claims. Similarly, hallucinations may sound factual but are inventions from the AI's internal language model probabilities, not from real sources.

Why AI Tools Like ChatGPT and Google Gemini Struggle with Citations

ChatGPT (by OpenAI) and Google's Gemini are powerful, large language models (LLMs) trained on vast text corpora. Their architecture generates fluent language by predicting word sequences rather than retrieving exact source documents. This difference leads to several challenges:

1. **No Built-in Source Indexing:** Most LLMs don't track precisely which training documents influenced specific output tokens.
2. **Knowledge Cutoff Dates:** ChatGPT's knowledge ends as of 2021 or early 2023 (depending on version), limiting current factuality.
3. **Generative vs. Retrieval Models:** Pure generative AI doesn't cite; retrieval-augmented generation systems (RAG) can, but require architecture integration.

Newer iterations of Gemini reportedly integrate more retrieval features, but as of now, requiring machine-sourced citations remains an evolving challenge.



Practical Strategies to Compel AI to Provide Sources

Despite limitations, here are some practical approaches when working with ChatGPT, Google Gemini, or similar tools:

1. Explicitly *Request Citations* or *Source Links* in Your Prompt

Always include instructions like:

"Please provide sources or links for all factual statements you cite."

For example, "Summarize the current state of Canadian private equity funds. Please quote exact sentences and provide the URL or document name for each fact."

The response may sometimes include footnotes or references, but beware that chatbots can fabricate plausible URLs and citations — so verification remains necessary.

2. Use Models or APIs with Built-in Retrieval Augmentation

Some AI systems combine generation with document retrieval, e.g., retrieval-augmented generation (RAG) frameworks. While ChatGPT itself is primarily generative, certain enterprise or research versions provide integrated citations by searching indexed databases live.

Google Gemini's planned enhancements focus heavily on multimodal reasoning and retrieving context, but as of yet, implementation details on citation transparency are limited. Keep an eye on product updates for such features.

3. Quote Extraction: Ask the AI to Extract Verbatim Passages

Prompt the tool specifically to:

- Include direct quotes from source documents
- Use quotation marks around the exact text
- Identify where the text is drawn from (e.g., *"According to the 2023 XYZ Private Equity Report, '...'"*)

This helps differentiate summary narrative from actual source material and facilitates manual cross-checking.

4. Verify Against Original Documents or Multiple Independent Sources

Family office teams should never rely solely on AI outputs without independent verification:

- Always cross-check summarized points against original managers' pitch decks, fund reports, or third-party databases.
- When possible, source data from reputable subscription platforms or regulatory filings rather than AI-only synthesis.

5. Maintain an "AI-Citation Log" as Part of Data Governance

From my experience, keeping an ongoing log of AI outputs alongside prompts and any references provided is invaluable. This document acts as an internal audit trail, answering the question I always ask before deploying a new tool: *"What would we show an auditor?"*

Special Considerations for Private Markets

Private markets represent an extreme use case due to scarce, non-standardized data and less public disclosure. AI hallucinations here can introduce material errors:

- Manager track records may be unavailable or inconsistent
- Valuation methodologies differ and are often not public
- Fund terms and structures have nuance hard to capture from text alone

This means AI tools need careful contextualisation; treat outputs as hypothesis generation rather than final fact. Prioritize traditional human due diligence and trusted data sources.

Summary Table: Requesting and Handling AI Sources

Action Purpose Limitations Practical Tip Request citations explicitly in prompt Push AI to indicate info origin AI might fabricate or omit Phrase requests clearly and specifically Use retrieval-augmented AI models Combine live data search with generation Not always available; may require subscription Explore API options and integrations Extract verbatim quotes Distinguish source text from narrative Quoted text quality depends on training data Ask for quotations with source ID Cross-verify with documents and multiple sources Detect errors, reduce risk Time-consuming but necessary Integrate with existing workflows Keep AI-citation logs Enables audits, tracking Requires disciplined process Make it standard team practice

Final Thoughts: Never Fully Trust AI Without Traceability

The progress in AI tools like ChatGPT and Google Gemini unlocks incredible potential for efficiency in family office investment research — but the downside risk of inaccurate, hallucinated, or unattributed information is real. Until generative models natively and reliably provide full transparent citations, the best defence is a combination of clear prompting, selective canadianfamilyoffices.com tool choice, and rigorous verification.

Demand your AI tools show where their information is coming from — and always ask “*What would we show an auditor?*” before incorporating AI insights into critical decision processes.

About the Author

With over 12 years of experience in UHNW multi-family offices across Canada, specializing in manager due diligence, reporting, and data governance, I focus on integrating technology while mitigating operational risk. I keep a running plain-text log titled “things AI got wrong” to sharpen scepticism and promote best practices.