

Generative AI tools such as **ChatGPT** and **Google Gemini** have transformed how information is processed, analysed, and presented. Their ability to generate natural language text, summarise complex documents, and parse data at scale is particularly appealing to ultra-high-net-worth (UHNW) family offices navigating vast rivers of financial data and manager reports.

However, these tools are not oracles. AI hallucinations—plausible but incorrect or fabricated outputs—pose risks that demand deliberate *human-in-the-loop* oversight, especially when it comes to investment decision-making, private markets due diligence, and sensitive reporting. This blog discusses what true human oversight <https://canadianfamilyoffices.com/special-reports/technology-sr/ai-has-the-makings-of-a-powerful-tool-for-family-offices-but-frequent-and-highly-convincing-hallucinations-are-proving-troublesome/> entails, the risks of overreliance on generative AI, and best practices for **review and approval** to ensure **accountability** and data integrity.

Understanding AI Hallucinations and Their Risks

One of the most insidious challenges with generative AI models like ChatGPT and Google Gemini is their propensity to produce *hallucinations*—content that is grammatically correct, contextually relevant, and highly persuasive but factually inaccurate or entirely fabricated.

- **Persuasiveness Breeds Trust:** These AI outputs often come packaged in professional language that sounds authoritative to non-experts. This makes it dangerously easy to mistake AI fabrications for facts.
- **Risk Amplified in Family Office Context:** The stakes are high. UHNW family offices rely on detailed, highly accurate data for manager due diligence, portfolio analysis, and regulatory compliance. Errors introduced by hallucinations can lead to poor investment decisions and loss of fiduciary trust.

For example, an AI summarizing a manager's quarterly report might invent a non-existent fund performance figure or misattribute risk metrics. Unlike public equity data with transparent pricing and multiple sources, *private markets data scarcity* compounds error risk because there are fewer signals to cross-check AI outputs.



Family Office Use Cases: Where Human Oversight is Critical

Generative AI is not without tremendous utility in family office operations, but it must be harnessed responsibly with human oversight:

1. Preliminary Research

Tools like ChatGPT and Google Gemini can rapidly parse large volumes of unstructured text (e.g., manager commentaries, market commentary) to produce initial summaries for analysts. However, these should be viewed as starting points—not final analyses—to be confirmed against source documents.

2. Parsing Complex Reports

AI can extract key sections and highlight relevant KPIs in dense manager reports, saving time for investment teams. Yet, given the subtleties in private placement memorandums or waterfall calculations, a skilled human reviewer must validate AI interpretations before relying on them for decisions.

3. Summaries and Briefings

Want to know something interesting? generating internal briefing notes or board-level summaries is a natural fit, but these outputs require rigorous fact-checking and contextual adjustments by subject matter experts to avoid perpetuating inaccuracies or glossing over critical nuances.

In each case, the generative AI serves as an assistant speeding data processing, *not* as an autonomous decision-maker.

Faithfulness Vs. Factual Hallucinations

When examining AI outputs, it is essential to distinguish two concepts:

Term	Meaning	Implications
Faithfulness	The extent to which generated content accurately reflects and is consistent with the source data. High faithfulness means summaries and extractions maintain the intent, context, and factual content of original documents.	
Factual Hallucinations	Fabricated facts or details that appear real but have no basis in source material or reality. Can lead to misguided investment decisions and compliance risk if treated as truth.	

Family offices must prioritise tools and workflows that maximise faithfulness and implement safeguards to detect and correct hallucinations before outputs are disseminated.

What Does Effective Human-in-the-Loop Oversight Look Like?

Human-in-the-loop (HITL) means human expertise guides, supervises, and intervenes in the AI's workstreams, especially for high-impact tasks. Here are key facets of effective HITL in generative AI for family offices:

1. Review and Approval Protocols:

- Every AI-generated product (research summary, report extraction, briefing note) must be reviewed by a qualified human analyst before circulation.
- Reviewers cross-check AI outputs against originals for accuracy, scope, and subtlety.
- Discrepancies are flagged and corrected prior to use.

2. Accountability and Audit Trails:

- Maintain logs of AI outputs, human reviews, and final approvals to support auditability.

- Document versions clearly so it is apparent when and how AI inputs were integrated and verified.

3. Tailored Training and Education:

- Educate family office staff on AI strengths, limitations, and common hallucination modes.
- Train reviewers to look for typical red flags and exercise scepticism.

4. Selective Use Cases:

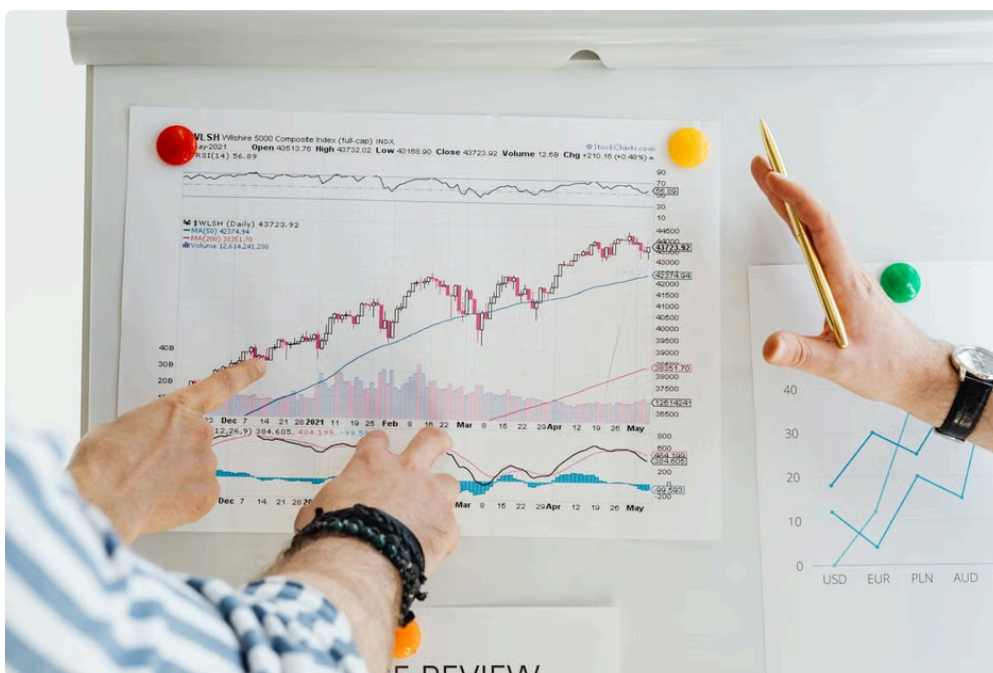
- Restrict generative AI to non-final-stage, low-risk tasks unless explicit human sign-off is in place.
- Avoid automating high-stakes compliance or investment decisions solely based on AI outputs.

5. Incremental Integration:

- Start with AI as a supplementary tool, iteratively incorporate feedback from human reviewers to refine processes.
- Continuously monitor AI performance and error rates.

A Practical Example

Imagine a family office analyst uses Google Gemini to summarise a private equity fund quarterly report. The AI produces a neat summary highlighting NAV growth, key investments, and fund expenses.



With proper oversight, the analyst:

- Reviews the summary alongside the original report to confirm no metric is misstated or omitted.
- Checks for invented performance data or misstated timelines (a common hallucination pitfall in private markets).
- Annotates corrections or comments as needed, then shares the vetted summary with the chief investment officer.
- Logs the review process and stores the AI output and final summary version for audit purposes.

This workflow reduces risk, improves accuracy, and preserves accountability across the team.

Why Family Offices Must Insist on Human Oversight

Unlike mass-market or entertainment AI applications, family office use is inherently high-stakes with regulatory, reputational, and financial impact. Blind trust in AI outputs is irresponsible because:

- **Private markets data scarcity:** makes hallucinations harder to detect.
- **Complex document structures:** increase misinterpretation risk.
- **Investment fiduciary duties:** impose legal accountability on families and advisors.

By embedding human judgement as a mandatory verification step, family offices protect themselves from costly errors and uphold governance standards expected by auditors and regulators.

Final Thoughts

Generative AI tools like ChatGPT and Google Gemini offer promising efficiencies, but their use in UHNW family offices demands a cautious, transparent approach anchored by rigorous *human-in-the-loop* oversight.

Here's a story that illustrates this perfectly: wished they had known this beforehand.. Building review and approval workflows ensures that AI outputs are factual and faithful representations rather than persuasive fabrications. Accountability mechanisms and audit trails provide governance visibility, a non-negotiable when stewarding multi-million-dollar portfolios and private market investments.

Ultimately, human expertise remains the linchpin that turns powerful generative AI tools from potentially risky black boxes into trusted augmentations of family office intelligence.