In today's fast-evolving AI landscape, organizations increasingly rely on multiple large language models (LLMs) — such as GPT, Claude, Gemini, Grok, and Perplexity — to support critical decision-making. These interactions, when orchestrated effectively, hold immense potential for surfacing and managing decision risks. But to capitalize on this, you need a systematic way to extract, validate, and document risks directly from AI conversations.

Enter the "risk register from an AI conversation" — a dynamic artifact capturing decision risks, their rationale, validation status, and action items. Combined with tools like **Scribe** for efficient documentation, this approach brings structure and rigor to what could otherwise be a nebulous exercise.

Why Build a Risk Register from AI Conversations?

Traditional risk registers are static and rely heavily on manual input from subject matter experts. However, AI conversations offer a rich, structured stream of risk details, concerns, hypotheses, and evidence — encapsulated in natural language across multiple perspectives and models.

Creating a risk register directly from these conversations provides several benefits:

- **Real-time risk surfacing:** Identify risks as they emerge during dynamic decision discussions.
- **Multi-model validation:** Cross-check risk points across different LLMs to uncover hallucinations or divergent insights.
- **Pressure-testing decisions:** Orchestrate conversations to probe decision assumptions and mitigate blind spots.
- **Shared context:** Maintain a coherent narrative by tracking inputs and outputs from GPT, Claude, Gemini, Grok, and Perplexity in one place.
- **Continuous improvement:** Use recorded interactions for auditability and iterative refinement of risk management practice.

Key Themes for Crafting an AI Conversation-Based Risk Register

1. Multi-Model Validation Within One Conversation

Different AI models have distinct architectures, training data, and tokenization strategies, which affects how they interpret and generate text. Relying on a single model increases the odds of blind spots or hallucinated facts creeping into your risk assessment.. (my cat just knocked over my water)

Multi-model validation means orchestrating a conversation involving several LLMs — for instance, asking GPT to identify potential decision risks, then sending those risks to Claude or Gemini for independent verification or contradiction.

This approach acts as a form of "cross-pollination," where confirmation from multiple models dramatically improves confidence in identified risks, while discrepancies flag areas needing deeper human review.

2. Pressure-Testing Decisions via Orchestration Modes

When you're orchestrating conversations across models, you can simulate multiple "modes" of reasoning:

- **Devil's advocate:** Force a model to argue against a proposed decision by emphasizing worst-case scenarios or overlooked risk factors.

- **Scenario analysis:** Ask models to envision a range of future states under different assumptions, highlighting decision sensitivities.
- **Consensus building:** Synthesize outputs from all models to identify common risk themes and conflicting views.

This pressure-testing ensures the decision maker is exposed to a fuller risk landscape rather than a narrow one-dimensional view.

3. Hallucination Detection Through Cross-Checking

Hallucinations — AI outputs that are plausible but factually incorrect or fabricated — are a notorious risk. By feeding the same question or risk statement to multiple models and comparing answers, contradictions become immediately visible.

For example, if GPT claims that a regulation applies in one jurisdiction but Claude disagrees, that inconsistency is a red flag warranting further investigation.

Cross-checking also involves fact-based validation — requesting source citations and dates, then manually or automatically verifying those references. Last month, I was working with a client who wished they had known this beforehand.. Remember, blindly trusting one model's "trust us" accuracy claims can lead to unsafe decisions.

4. Keeping Shared Context Across GPT, Claude, Gemini, Grok, and Perplexity

Context is king in complex conversations. AI models have token limits and can lose track of earlier discussion threads. A robust orchestration framework maintains:

- **Shared state:** Tracking prompts, questions, model outputs, and generated risks in a centralized repository.
- **Context stitching:** Feeding summarized prior interactions as input to subsequent model queries to maintain topic continuity.
- **Versioning:** Log changes as risk descriptions evolve through feedback cycles.

These practices ensure that each subsequent model iteration builds meaningfully on prior insights rather than starting from scratch or reiterating the same points.

Step-by-Step Guide: Building a Risk Register from an AI Conversation

1. **Define Decision Scope and Objectives** Before launching AI queries, clearly articulate the decision or project scope. Document the core question and desired outcomes. This sets boundaries for risk extraction.
2. **Initiate Initial Risk Identification with GPT** Prompt GPT to list potential decision risks based on the scope. Ask for explanations and underlying assumptions. Extract and format these into preliminary risk register entries.
3. **Cross-Validate Risks Using Claude and Gemini** Send each risk and explanation to Claude and Gemini with queries such as:
    - "Do you agree this is a risk? Why or why not?"
    - "What additional risks might be related or overlooked?"

    Annotate discrepancies and new insights in the risk register.
4. **Apply Pressure-Testing via Orchestration Modes** Use Grok or Perplexity to:
    - Devil's advocate each risk or decision point.

- Run scenario analyses (best case, worst case, most likely).
- Request summaries highlighting consensus and dissent.

Add new entries or update existing risks to reflect these explorations.

5. **Detect and Flag Hallucinations** Cross-check factual claims across all models. Where disagreements arise, note for human validation or further research.
6. **Keep Context Updated in a Central Repository** Use a tool like **Scribe** to document the entire conversation thread, from initial prompts to final outputs. Organize notes, timestamps, model names, and risk status fields (e.g., identified, validated, mitigated).
7. **Review and Iterate with Human Experts** Share the risk register with stakeholders for feedback. Incorporate corrections, additional context, or mitigation plans. Use AI to support the update cycle, not replace it.
8. **Establish Continuous Monitoring** Enable follow-up conversations to update the risk register as new information emerges or assumptions shift.

Example: Risk Register Table from Multi-Model AI Conversation

Risk ID Risk Description Source Model(s) Validation Status Notes / Discrepancies Mitigation Actions R1 Compliance risk due to ambiguous data privacy regulations in EU post-2023. GPT, Claude, Gemini Partially validated GPT and Claude agree; Gemini questions applicability to marketing data. Flagged for human legal review. Consult legal team; update data handling policies accordingly. R2 Operational risk from AI hallucination of competitor pricing. GPT, Grok, Perplexity Flagged as hallucination Grok and Perplexity disagree with GPT's claim on competitor pricing. Source citations absent. Exclude this data until verified; request up-to-date market report. R3 Technology risk due to vendor lock-in with proprietary AI APIs. All models Validated Consensus reached; scenario analysis predicts high switching costs. Evaluate multi-vendor strategy; negotiate contract terms.

What Would Change My Mind?

I am advocating for building risk registers directly from AI conversations enhanced by multi-model orchestration. However, this approach presumes:

- All team members trust and understand AI-generated insights.
- Models are sufficiently dissimilar to yield meaningful validation.
- Organizations have the discipline and tooling (like Scribe) to maintain shared context robustly.

If future evidence emerges that AI-driven risk registers increase noise or false positives instead of clarity, or if orchestration complexity outweighs benefits, I would reconsider the recommended approach.

Avoiding the "Five Tabs in a Trench Coat" Trap

Be wary of AI platforms that simply layer multiple LLMs behind a single UI without true https://www.launchboard.dev/launch/suprmind-1328 orchestration or context management — effectively "five tabs in a trench coat." Without deep integration, multi-model validation becomes little more than a manual copy-paste exercise, defeating the purpose of automation and increasing cognitive load.

Effective risk registers from AI conversations require purpose-built orchestration, transparency in model usage (name your models!), and real-time context stitching to extract meaningful, actionable decision risks.

Conclusion

Want to know something interesting? creating a risk register from an ai conversation is both an art and a science. By harnessing the complementary strengths of diverse AI models, pressure-testing decision assumptions, vigilantly detecting hallucinations, and maintaining unified context with tools like Scribe, organizations can unlock new levels of rigor in decision risk management.

As AI adoption accelerates, this approach empowers teams to convert dynamic, natural language insights into disciplined risk registers — a crucial step to make AI-driven decision-making safer, more transparent, and ultimately more trustworthy.