The AI Investment Decision Framework

Should You Actually Use AI for This?

Many businesses burn significant capital on custom LLM deployments that result in zero efficiency gains, while missing simple, low-cost automated workflows that can save thousands of hours in manual data entry.

The difference is never the technology. It's the target.

Most businesses look at AI and ask, "What can this do?" That is a guaranteed way to waste money. The correct question to ask is, "What painful, expensive, repetitive bottleneck do we already have that a machine can execute?"

This framework is designed to force hard thinking. No 1-10 wishy-washy scales. No buzzwords. Just a brutal filter to ensure AI investments are directed where they will actually yield a return.

Prerequisite: The 6-Month Rule

If the task has not been done manually for 6 months, it should not be automated.

Why: AI is not a magic wand for broken processes. It is a strict amplifier. Applying AI to a disorganized, undefined process yields messy, unpredictable results at scale. Before introducing an LLM, human operators need to have executed the task long enough to define the edge cases, establish standard operating procedures, and define what "good" looks like.

If launching a brand new service line today, execute it manually. Revisit AI in six months.

The 4-Question Scoring Framework

To proceed with an AI project, the task must score four hard "YES" answers. One "NO" means the project dies or requires a fundamental redesign.

Question 1: TIME

"Does this consume >3 hours/week of actual time?"

- **The Reality Check:** Human estimation of time is often inaccurate. Before answering, require the use of a time-tracking tool (like Toggl or Clockify) on this specific task for one week.
- **The Trap:** There is a tendency to automate disliked tasks, even if they only take 15 minutes a week. Setting up, testing, and maintaining an AI workflow for a 15-minute task will cost more time and money than executing the task manually for the next five years.
- **RED FLAG:** "It only takes 20 minutes but it's really annoying." Kill the project.

Question 2: REPEATABILITY

"Is the task 80%+ identical each time?"

- **The Reality Check:** Current AI models excel at patterned execution, not novel strategy. If a bulletproof, 10-step checklist can be written for this task that an entry-level worker could follow perfectly without asking questions, AI can likely execute it.
- **The Creative Work Trap:** AI should not be used to invent a go-to-market strategy or design a brand-new product from scratch. It should be used to extract 50 invoice numbers from 50 wildly different PDF formats and output them to a CSV. High variance in input format is acceptable; high variance in the required *logic* is fatal.
- **RED FLAG:** "Every client situation is completely unique and requires a custom approach." Kill the project.

Question 3: ERROR COST

"If the AI gets it wrong, can it be caught in <10 minutes and fixed for <$50?"

- **The Reality Check:** Generative AI hallucinates and presents false information confidently. A production workflow must assume a standard failure rate. Therefore, a "human in the loop" is required to catch errors before they impact the customer or the business's finances.
- **Documented Failures:** * *The $1 Chevy:* A car dealership deployed an unmonitored AI chatbot that agreed to sell a customer a 2024 Chevy Tahoe for $1, resulting in a public relations and legal headache.
    - *The Phantom Precedent:* A lawyer used ChatGPT to write a brief without verifying the citations. ChatGPT invented fake court cases, and the lawyer was officially sanctioned by the judge.
- **RED FLAG:** "The output goes directly to the client/live database without human review."

Question 4: DATA SAFETY

"Can this be executed without feeding proprietary customer, employee, or financial data into a public model?"

- **The Reality Check:** Pasting raw financial data or client lists into the free, consumer version of standard LLMs typically violates NDAs and privacy policies, as that data may be used to train future iterations of the model.
- **RED FLAG:** "We'll just copy-paste the customer database into the chat window." Kill the project.

The 1-Page Vendor Security Assessment (If using a third-party AI tool):

Before procuring an AI SaaS tool or wrapper, require the vendor to answer the following. Any hesitation is a reason to pause the procurement.

1. **Zero Data Retention / Non-Training:** "Do you explicitly state in your enterprise MSA that our data is NEVER used to train your foundational models?" (Must be YES).
2. **Data Residency:** "Where exactly is our data hosted and processed?"
3. **Compliance:** "Are you SOC2 Type II certified?"
4. **Third-Party APIs:** "Are you passing our data to external APIs (OpenAI, Anthropic)? If yes, are you utilizing zero-retention Enterprise APIs?"

Tool Selection Matrix

Select the correct tool based on its verified strengths, rather than generalized hype.

Tool	When to Use It	When to Avoid It
Claude (Anthropic)	Writing that requires a natural tone, coding, and parsing massive documents. It currently has a leading "context window" and follows complex formatting instructions highly accurately.	Live web browsing or deep ecosystem integrations.
ChatGPT (OpenAI)	General reasoning, advanced data analysis/Python execution, voice interactions, and custom GPTs for structured internal team use.	Writing final marketing copy (it tends to output repetitive, robotic structures).
Gemini (Google)	Heavy reliance on the Google Workspace ecosystem. Excellent for summarizing YouTube content or parsing data directly from Gmail/Drive integrations.	Highly complex coding tasks (competitors currently edge it out).
Custom API / Make.com	The process passed the 4-question framework and needs to run in the background 100 times a day automatically.	The base prompt has not yet been proven to work manually.
Nothing	The task requires human empathy, strategic intuition, or falls under the "kill list."	N/A

The "Kill List": 5 AI Projects That Routinely Fail

Avoid investing resources into these commonly pitched, high-failure-rate initiatives.

1. **Fully Autonomous Customer Support for High-Ticket Items:** Clients paying premium rates expect human interaction when issues arise, not an automated widget.

2. **"The Everything Chatbot":** Connecting an AI to an entire, unstructured company cloud drive. It will inevitably hallucinate, surface outdated policies, or expose restricted internal documents. Start with a tightly scoped, single-purpose database.

3. **The Unsupervised Sales Outreach Cannon:** Automating LinkedIn or email to send thousands of hyper-personalized messages unreviewed. The AI will eventually misinterpret a prospect's bio, sending inappropriate messages and damaging domain reputation.

4. **Replacing Copywriters Entirely:** AI generates structure efficiently but struggles with distinct brand voice. Removing human writers entirely results in generic, highly detectable AI content. AI is best used to generate volume, while humans refine and select.

5. **Complex Multi-Agent Workflows for Simple Forms:** Designing an ecosystem of multiple AI agents to qualify a standard lead. A standard web form or Typeform is faster, cheaper, and more reliable.

The AI Project Scoring Sheet (Printable)

Project Name: _____

Proposed By: _____

Date: _____

Prerequisite Check

[ ] **The 6-Month Rule:** This exact process has been executed manually for at least 6 months. The edge cases are known and documented.

If unchecked, STOP. Project is killed.

The 4-Question Framework

(All questions must be answered YES to proceed).

1. TIME: Does this task consume >3 hours of actual, tracked time per week?

[ ] YES

[ ] NO *(Red Flag: Waste of setup time)*

2. REPEATABILITY: Is the underlying logic of this task 80%+ identical every single time?

[ ] YES

[ ] NO *(Red Flag: Requires novel, strategic human thought)*

3. ERROR COST: If the AI hallucinates, can a human catch it in <10 minutes and fix it for <$50 without client impact?

[ ] YES

[ ] NO *(Red Flag: Unsupervised outputs or catastrophic liability)*

4. DATA SAFETY: Can we execute this without putting proprietary/PII data into a public model that trains on our inputs?

[ ] YES

[ ] NO *(Red Flag: Massive security/NDA violation)*

Final Verdict

Score (Count of YES answers): ____ / 4

- **4/4 YES:** Approved for prototyping.
- **Anything less than 4:** KILLED. Back to the drawing board.

Next Step (If Approved):

Tool to be used for initial prototype: _____

Name of the "Human in the Loop" who will review outputs: _____