

How to Measure the ROI of AI Projects

From AI Experimentation to Measurable Business Value

KIMBERLY WIETHOFF, MBA, PMP, PMI-ACP

AI GOVERNANCE · PMO · BENEFITS REALIZATION

Managing Projects The Agile Way

#AI #ArtificialIntelligence #AIROI #EnterpriseAI #AITransformation
#ProgramManagement #ProjectManagement #PMO #PMOLeadership
#BusinessValue #BenefitsRealization #AIGovernance #AgenticAI
#DigitalTransformation #AILEadership #ManagingProjectsTheAgileWay



The Question Every Executive Is Asking

The Technology Is Advancing Rapidly

Organizations are investing in copilots, intelligent automation, predictive analytics, generative AI, machine learning, and increasingly, AI agents capable of performing work — not just assisting with it.

During a pilot, success feels easy. A model produces an impressive result. An agent completes a workflow in seconds. Everyone sees the potential.

But the Harder Question Is Now:

Are these investments actually creating measurable business value?

Once organizations move from experimentation to enterprise deployment, executives eventually ask: **"What value did we actually receive from the investment?"**

That is where many AI initiatives struggle. It is relatively easy to demonstrate that AI *works*. It is much harder to prove that AI improved the business enough to justify the investment.



Benefits realization needs to begin before the project starts — not after go-live.



AI Success $\neq$ Business Success

Technical and adoption metrics can look impressive — and still tell you nothing about ROI.

Technical Performance

95% model accuracy. Chatbot answered 50,000 questions. 10,000 agent tasks completed.

Tells us whether the technology **works**.

Adoption Metrics

85% of licenses activated. 100,000 prompts generated. Pilot launched on schedule.

Tells us whether people are **using it**.

Business Outcomes

Lower costs. Faster processes. Higher productivity. Increased revenue. Fewer errors.

Tells us whether the investment **mattered**.

Executives are not investing in prompts, models, agents, tokens, or licenses. They are investing because they expect something in the organization to improve. **The PMO's job is to connect the technology investment to those outcomes.**

Start With the Business Problem, Not the AI Solution

One of the most common mistakes in AI programs: beginning with the technology instead of the outcome.

✗ Solution-First Framing

- "We need an AI assistant."
- "We should build an AI agent."
- "We need a generative AI strategy."

These describe potential solutions. They do not describe business problems.

✓ Outcome-First Framing

- "Reduce employee search time by 30%."
- "Cut reporting prep from 6 hours to 2 hours weekly."
- "Resolve 40% of routine inquiries without human intervention."

Now you have something measurable. The AI becomes the **solution**. The business outcome becomes the **goal**.

That distinction fundamentally changes how the project is governed — and how success is defined from day one.

Establish the Baseline Before Implementing AI

You cannot prove improvement if you do not know where you started. Before implementing AI, capture the current state of the process — even if it is messy and inconsistent.

What to Measure

- Hours required to complete the process
- Cost per transaction
- Cycle time and error rate
- Customer and employee satisfaction
- Rework, escalation, and defect rates
- Revenue conversion and compliance incidents

Don't Wait for a Perfect Baseline

Different teams may perform the same work differently. Some employees complete a task in 20 minutes; others take an hour. Establish a **credible baseline range** instead.

Measure across multiple employees, teams, transaction types, time periods, and performance levels. Then document the variation.

Example: "Current reporting preparation time: 4–7 hours per week, median 5.5 hours."

That is far more useful than pretending the organization has a perfectly standardized process. The goal is credible evidence — not mathematical perfection.

Define the Value Hypothesis

Before funding any AI initiative, document a simple, structured value hypothesis that gives the steering committee something concrete to evaluate.

If we implement **[AI capability]**, we expect **[business metric]** to improve from **[baseline]** to **[target]**, creating approximately **[business value]** within **[time period]**.

Example in Practice

If we implement AI-assisted project reporting, we expect reporting preparation time to decrease from an average of five hours per program manager per week to two hours — creating approximately **3,000 hours of additional annual delivery capacity**.

The Problem

Defined and scoped

The Baseline

Measured and documented

The Target

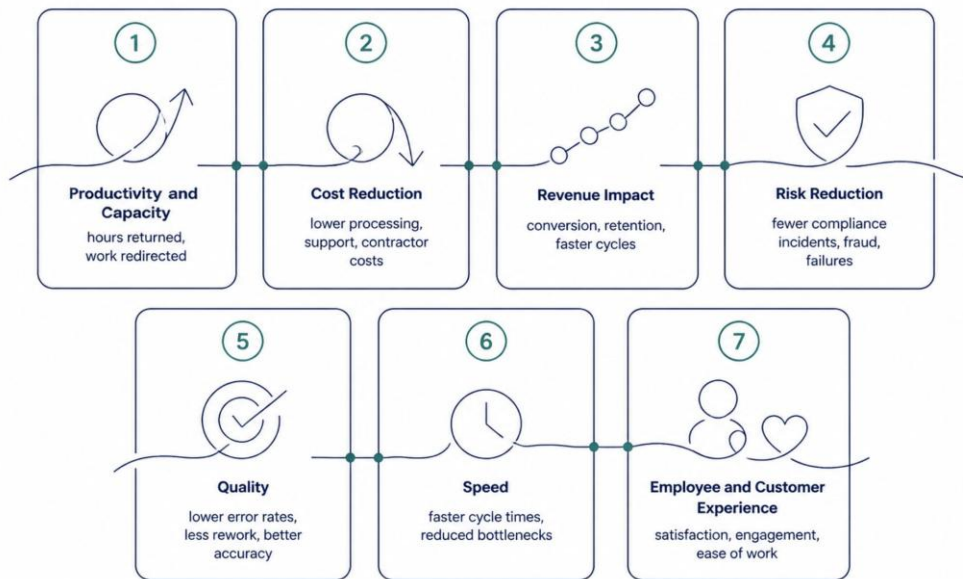
Specific and time-bound

The Value

Estimated and governed

Seven Dimensions of AI Value

AI value does not always fit neatly into one financial metric. Evaluate investments across multiple dimensions to build a complete and credible business case. The traditional ROI formula ($\text{ROI} = (\text{Benefit} - \text{Cost}) \div \text{Cost} \times 100$) remains the foundation. But a \$500K investment yielding \$900K in measurable benefits (80% ROI) only tells part of the story. Capture value across all seven dimensions for a complete picture.



Productivity: Capacity Created vs. Money Saved

The Calculation

Suppose 20 program managers each spend five hours per week preparing reports — **100 hours per week**. AI reduces that to two hours per person.

60 hours are returned to the organization every week. Across 52 weeks: **3,120 hours of additional capacity annually.**

What That Capacity Enables

- Risk and dependency management
- Stakeholder engagement and coaching
- Strategic planning and benefits realization
- Vendor management and delivery improvement

A Critical Distinction

Time saved is **not** automatically money saved.

If no salaries change and no positions are eliminated, the organization did not "save" the salary equivalent of those hours.

What it created was **capacity** — and that capacity is genuinely valuable when redirected toward higher-value work.

- ❏ Claiming millions in "savings" based on hours never removed from the cost structure is one of the fastest ways to undermine an AI business case.

Risk Reduction: Measuring What Didn't Happen

Some of the most valuable AI systems create value by preventing something — fraud, compliance violations, operational failures, safety incidents. This can be difficult to quantify, but there is a practical approach.

The Framework

Expected Risk Exposure = Probability of Event × Financial Impact

An operational failure with a 20% annual probability and \$5M impact carries **\$1M in expected annual exposure**.

The AI Improvement

If AI-enabled monitoring reduces that probability from 20% to 5%, expected exposure drops to **\$250,000**.

Estimated risk reduction: **\$750,000** — not a check received, but a material reduction in financial exposure.

Why It Matters

For highly regulated or operationally complex organizations, risk reduction can be AI's most important form of value — often exceeding the direct productivity or cost savings from the same investment.



Don't Forget the Full Cost of AI

One of the biggest mistakes in AI business cases is underestimating the denominator. The cost of AI is not simply the software license — and production economics are very different from pilot economics.

Infrastructure & Platform

- AI platform and API costs
- Cloud infrastructure and data storage
- Model consumption at scale

Governance & Compliance

- Cybersecurity and AI governance
- Testing, validation, and compliance
- Model monitoring and maintenance

People & Change

- Training and change management
- Human review and support operations
- Integration and data preparation

⚠ A pilot might serve 50 users. Enterprise deployment might serve 10,000. Production changes the economics — always evaluate total cost of ownership, not just the cost of experimentation.

AI Agents Change the ROI Equation

Traditional AI

Assists employees. Answers questions. Generates drafts. Reduces search time. Humans retain control of every workflow step.

Value measured in: time saved, quality improved, user satisfaction.

Agentic AI

Performs portions of workflows autonomously. Completes tasks, not just recommendations. Organizations can now measure entirely new metrics.

- Percentage of workflow completed autonomously
- Human interventions per transaction
- Cost per autonomous transaction
- Agent failure and escalation rates

What Strong Agent Reporting Looks Like

"The agent successfully completed 72% of eligible service requests without human intervention, reduced average handling time by 63%, lowered cost per transaction by 28%, and maintained customer satisfaction above the established threshold."

That tells executives what the agent accomplished **for the business** — not just how many tasks it ran.

Build an AI Value Scorecard

Every significant AI initiative should have a value scorecard that gives leadership a real-time view of business outcomes — not just delivery milestones.

Metric	Baseline	Target	Current	Trend
Process cycle time	5 days	2 days	2.4 days	 Improving
Error rate	12%	6%	7%	 Improving
User adoption	0%	80%	76%	 Improving
Cost per transaction	\$18	\$12	\$13	 Improving
Customer satisfaction	84%	90%	91%	 Above Target
Annual value realized	\$0	\$1.5M	\$1.1M forecast	 At Risk

The scorecard shifts the governance conversation from *"How is the AI project going?"* to **"Are we achieving the outcomes we funded?"** — a fundamentally more strategic question.

Leading vs. Lagging Indicators

A mature AI value framework tracks both. Relying only on one type leaves your governance blind to half the story.

Leading Indicators

Tell us whether the organization is moving **toward** value.
Monitor early and often.

- Adoption and usage rates
- Model accuracy and agent success rates
- Process compliance and training completion
- Employee engagement with the tool

Leading indicators tell us what **may happen**.



Lagging Indicators

Tell us whether value was **actually realized**. Measured at defined checkpoints.

- Cost savings and revenue increase
- Cycle-time and error reduction
- Customer and employee satisfaction
- Productivity improvement and risk reduction

Lagging indicators tell us what **did happen**.

Track value across a defined cadence: **Pilot** → **30 Days** → **90 Days** → **6 Months** → **12 Months**. At each checkpoint, assess whether to continue, adjust, scale, or stop.

When AI Doesn't Deliver Expected ROI

Not every AI initiative will succeed — and continuing to invest because of sunk cost is the bigger failure. When value is not materializing, investigate systematically.

1 Diagnose the Root Cause

Is adoption too low? Is the data poor? Are users bypassing the solution? Is the model underperforming? Are AI outputs requiring too much human correction? Were assumptions unrealistic?

2 Identify the Right Intervention

Sometimes the answer is to improve the solution, redesign the workflow, change the model, or strengthen adoption. Each situation requires a different response.

3 Be Willing to Stop

Stopping a low-value AI initiative is not failure — it can be excellent portfolio management. The goal is to direct investment toward initiatives creating the greatest strategic value, not to keep every AI project alive.

The PMO's Role: From Project Status to Value Status

The PMO Can Own Benefits Realization - AI teams focus on whether the technology works. Business teams focus on whether the process works. Finance focuses on the investment. Cybersecurity focuses on risk. Executives focus on strategy. **The PMO can connect all of them** — making benefits realization one of the most important PMO opportunities in enterprise AI transformation.

What the PMO Should Establish

- Value hypotheses and measurement frameworks
- Benefits owners and governance checkpoints
- Executive scorecards and scale/stop decisions
- Business outcome definitions and baseline requirements

❏ Go-live is a delivery milestone — not proof of business value. Benefits realization continues until the organization can answer: "Did we receive the value we expected?"

The New Executive Dashboard

Expected Annual Value \$2.4M

Value Realized to Date \$1.1M

Productivity Improvement 31%

Autonomous Completion 68%

ROI Forecast
137% — On Track

Leadership is no longer reviewing whether projects were delivered. They are evaluating whether **investments are successful**.

The Organizations That Will Win With AI

The most successful organizations will not necessarily be the ones with the most AI pilots, copilots, models, or agents.

Start With the Problem

Define the business outcome first. Let AI be the solution, not the starting point.

Measure What Matters

Establish baselines, define value hypotheses, and track business outcomes — not just adoption or accuracy.

Account for Full Cost

Evaluate total cost of ownership across platform, governance, people, and change — especially at production scale.

Govern the Portfolio

Scale what works. Improve what's struggling. Stop what isn't creating value. Direct investment toward the greatest strategic return.

They will be the organizations that know which AI investments are creating value, which ones are not — and where they should invest next.

By Kimberly Wiethoff, MBA, PMP, PMI-ACP — Senior Program Manager and AI Transformation Leader · Managing Projects the Agile Way