AI Decision Effectiveness Scorecard

Rate where AI sits in your workflow. For each cell: read the question, then circle one rating.

IRDA: Information !' Recommendation !' Decision !' Action | mile-hi.ai/frameworks/irda

DECISION MATURITY	I Information	R Recommendation	D Decision	A Action
I Individual	Can one operator find the facts they need, when they need them? ☐ Not yet ☐ Starting ☐ Working ☐ Strong	Does AI suggest the right next step for this person's context? ☐ Not yet ☐ Starting ☐ Working ☐ Strong	Is the operator deciding — or has AI quietly taken over? ☐ Not yet ☐ Starting ☐ Working ☐ Strong	Does the recommended action actually get taken and tracked? ☐ Not yet ☐ Starting ☐ Working ☐ Strong
II Team	Can the team share a single AI-structured source of truth? ☐ Not yet ☐ Starting ☐ Working ☐ Strong	Do AI recommendations land consistently across the team? ☐ Not yet ☐ Starting ☐ Working ☐ Strong	Are team decisions more coherent with AI in the loop? ☐ Not yet ☐ Starting ☐ Working ☐ Strong	Do team-level actions compound, not just individual ones? ☐ Not yet ☐ Starting ☐ Working ☐ Strong
III Organization	Can the org surface cross-functional patterns via AI? ☐ Not yet ☐ Starting ☐ Working ☐ Strong	Does AI enable decisions the org couldn't previously take? ☐ Not yet ☐ Starting ☐ Working ☐ Strong	Is decision authority clear when AI spans departments? ☐ Not yet ☐ Starting ☐ Working ☐ Strong	Are org-wide actions auditable and tracked end-to-end? ☐ Not yet ☐ Starting ☐ Working ☐ Strong
IV Ecosystem	Does AI structure market-level data (MRFs, benchmarks)? ☐ Not yet ☐ Starting ☐ Working ☐ Strong	Can participants act on AI-structured comparisons? ☐ Not yet ☐ Starting ☐ Working ☐ Strong	Do ecosystem actors decide differently with this data? ☐ Not yet ☐ Starting ☐ Working ☐ Strong	Are ecosystem actions (procurement, pricing) transformed? ☐ Not yet ☐ Starting ☐ Working ☐ Strong

How to use: Circle one rating per cell. Cells where you marked "Not yet" or "Starting" at higher tiers (Org/Ecosystem) are your highest-leverage AI investment gaps.

How to Use the Scorecard

Each scorecard is for one AI project. A project can impact an individual's productivity, a team's productivity, an organization's productivity, or an entire ecosystem of partners and suppliers. Score the row that matches the tier your project primarily operates at.

1. Pick one AI project.

Don't score "the organization" — pick one specific AI initiative. "Our AI trading assistant" or "the enterprise risk aggregation model" is the right level of specificity. If you have five AI projects, fill five scorecards.

2. Identify the primary tier.

Which tier does this project primarily impact? An AI signal screener for one trader = Individual. A research synthesis tool shared by a trading desk = Team. A cross-divisional risk framework = Organization. A systemic-risk model shared across institutions = Ecosystem.

3. Score across all four IRDA stages.

For the primary tier row, read each cell's question and circle one rating. Then check the rows above and below — AI projects often have secondary effects at adjacent tiers.

4. Read the ratings: Not yet / Starting / Working / Strong.

Not yet = capability doesn't exist. Starting = early experiments underway. Working = used regularly and reliably. Strong = load-bearing infrastructure; other decisions depend on it.

5. Find your highest-leverage gap.

Look for the rightmost "Working" or "Strong" cell in a row. The cell immediately to its right (the next IRDA stage) is your highest-leverage investment — the AI below it works, but the next stage hasn't been built.

6. Compare and revisit.

Fill independently, compare with your team. Disagreements reveal misaligned assumptions. Revisit quarterly — monthly is noise, annually misses regressions.

Worked Examples — One Project Per Tier

Each example shows a real AI project, the tier it primarily impacts, and a completed scorecard row.

I. Individual — AI-Assisted Stock Trading for an Individual Trader

A trader uses an AI tool that monitors real-time market feeds, SEC filings, earnings transcripts, and social sentiment to surface trade-relevant signals. The AI flags anomalies, suggests entry/exit points based on technical and fundamental analysis, and tracks whether executed trades matched the AI's thesis.

	Information	Recommendation	Decision	Action
Individual	**Strong**	**Working**	**Starting**	**Not yet**

Reading: Information is "Strong" — the AI reliably aggregates multi-source signals (price feeds, filings, sentiment) faster than the trader can manually. Recommendation is "Working" — entry/exit suggestions are used daily and improve hit rate. Decision is "Starting" — the trader still overrides ~40% of suggestions based on gut feel; the AI doesn't model the trader's personal risk tolerance or portfolio constraints well enough to be trusted. Action is "Not yet" — no automated execution or post-trade tracking links back to the AI's original thesis. Next move: improve Decision by encoding the trader's risk parameters and position-sizing rules into the AI's recommendation context.

II. Team — AI Market Research for a Fixed-Income Trading Desk

A team of fixed-income traders uses AI to synthesize research across their segment — aggregating sell-side analyst reports, central bank communications, yield-curve models, and credit-spread patterns. The AI produces a daily briefing with trade recommendations ranked by conviction and risk-reward ratio for the desk's specific book.

	Information	Recommendation	Decision	Action
Team	**Strong**	**Working**	**Starting**	**Not yet**

Reading: Information is "Strong" — the team shares a single AI-structured research layer that replaces fragmented email-and-PDF workflows. Recommendation is "Working" — the daily briefing with ranked trades is used in the morning meeting and shapes ~60% of the desk's activity. Decision is "Starting" — senior traders still re-rank the AI's suggestions before committing capital; the model doesn't account for the desk's existing exposure and concentration limits. Action is "Not yet" — there's no systematic tracking of whether AI-recommended trades outperform trader-originated ones. Next move: integrate the desk's live position book into the AI's ranking engine so recommendations account for existing exposure.

III. Organization — AI Risk Evaluation Framework for a Financial Institution

A bank deploys an AI system that aggregates risk signals across all business lines — credit risk, market risk, operational risk, and liquidity risk. The model identifies correlated exposures that siloed risk teams miss, surfaces emerging concentration risks, and generates enterprise-wide stress-test scenarios that feed into capital-allocation decisions.

	Information	Recommendation	Decision	Action
Organization	**Working**	**Starting**	**Not yet**	**Not yet**

Reading: Information is "Working" — the AI reliably aggregates cross-divisional risk data that previously lived in separate systems. Recommendation is "Starting" — the model generates stress scenarios but its correlation assumptions are challenged by risk committees ~50% of the time. Decision and Action are both "Not yet" — capital-allocation decisions are still made in quarterly committee meetings using static reports, and there's no automated link between AI-identified concentrations and hedging workflows. Next move: improve Recommendation by validating AI correlation models against historical stress events before pushing for Decision-tier adoption.

IV. Ecosystem — AI Risk Evaluation Framework for a Financial Ecosystem

A consortium of banks, insurers, and regulators uses AI to evaluate systemic risk across the financial ecosystem. The model structures counterparty exposure data, cross-institution lending networks, and regulatory filings to surface contagion pathways — risks that no single institution can see from its own data alone.

	Information	Recommendation	Decision	Action
Ecosystem	**Starting**	**Not yet**	**Not yet**	**Not yet**

Reading: Information is "Starting" — data-sharing agreements exist but coverage is incomplete; only ~30% of counterparty exposure data flows into the shared model. Recommendation, Decision, and Action are all "Not yet" — no AI-driven systemic risk recommendations exist at ecosystem scale, no consortium-level decision framework consumes AI output, and regulatory response workflows remain manual. Next move: achieve "Working" at Information by expanding the data-sharing consortium and standardizing the exposure data format before attempting ecosystem-level Recommendations.

Frequently Asked Questions

Q: How often should I fill this out?

Quarterly. AI maturity shifts slowly — monthly is noise, annually is too slow to catch regressions. If you're mid-implementation on a specific project, a focused re-score at the 90-day mark is also useful.

Q: Should I fill one for the whole organization or per project?

Per project or per AI initiative. "The organization" is too broad — you'll average away the interesting signals. Pick one AI system and grade it specifically. If you run three AI projects, fill three scorecards.

Q: What if my project only touches Individual and Team tiers?

That's normal. Most AI deployments today live at Individual. The scorecard helps you see what "graduating" to Team or Organization would require — which is often a decision-architecture question, not a technology question. The empty rows aren't a failure; they're a map of what's next.

Q: What's the difference between "Working" and "Strong"?

"Working" means the capability exists and is used regularly. "Strong" means it's reliable enough that other decisions depend on it — it's load-bearing infrastructure, not an experiment. A recommendation engine that's "Working" gives useful suggestions. One that's "Strong" is what the team uses to set priorities.

Q: We scored "Not yet" on most Decision and Action cells. Is that bad?

It's typical, not bad. Most production AI lives at Information and Recommendation. The D and A columns are where the real organizational leverage is — and where most teams haven't arrived yet. That's the point of the scorecard: making the gap visible so you can invest deliberately.

Q: How does this relate to the IRDA framework?

IRDA (Information !’ Recommendation !’ Decision !’ Action) is the horizontal axis. Each column represents a stage where AI can sit in a workflow. The vertical axis is Decision Maturity — the scale at which decisions happen (Individual !’ Team !’ Organization !’ Ecosystem). The scorecard maps these two dimensions into a 4×4 assessment grid. Learn more at mile-hi.ai/frameworks/irda

Q: Can I share this with my team?

Yes — it's ungated and free. The most valuable use is filling it out independently, then comparing ratings as a team. Disagreements about ratings are where the real conversations happen — they surface misaligned assumptions about what AI is actually doing vs. what people think it's doing.

Q: What should I do after filling it out?

Look for the highest-tier cell where you marked "Not yet" or "Starting" that sits to the RIGHT of a "Working" cell in the same row. That's your highest-leverage gap: the AI capability below it works, but the next stage hasn't been built. That's where investment compounds fastest.