# Why AllmaGen

## The AI-Powered Decision And Automation Suite For High Accuracy With Low Risk

30 June 2025


### Summary

Smart companies investing in AI prioritize applying it to decision quality — not just execution efficiency. Why? Because it doesn't matter how efficiently AI helps you operate if you're steering the business in the wrong direction through poor decision-making. So it's no surprise that many decision intelligence (DI) technology providers have begun injecting AI technologies into their products. However, AI models and agents are probabilistic, As a result, they suffer from hallucinations, leading to inaccuracies and inconsistent results. And research shows their ability to reason correctly collapses when faced with the high complexity inherent in high-stakes enterprise decision-making.

To overcome this challenge, AllmaGen has created and field-tested an innovative hybrid approach (patent pending) overlaying a deterministic framework on embedded probabilistic AI technologies. It quantifies risks and recommends mitigating actions — boosting accuracy, preventing costly errors, and explaining and justifying the decisions it recommends, while also enabling automation.

Based on this framework, unlike other AI platform providers, AllmaGen has built solutions for specific industry functions: Project Manager, RFP Manager, and Research Author — with more on the way. And we've proven we can build new ones rapidly that can be trusted, thanks to the universality of our platform.

## Contents

# Overview

## High-Stakes Decision-Making Is The Most Valuable — And Dangerous — Use of AI

It doesn't matter how efficiently AI helps you operate if you're steering the business in the wrong direction through poor decision-making. That's why AI investments should prioritize decision quality over execution efficiency, to infuse intelligence into high-stakes processes (whether fully automated or human-in-the-loop). So it's no wonder many DI technology providers are injecting AI technologies into their products. The potential ROI is high for midsize to large enterprises, where IT struggles to keep up, many business applications still run on spreadsheets, and budget to hire  is scarce — especially in complex industries like construction, engineering, finance, and IT selection research.

But a dark cloud looms over AI-powered decision-making technologies because AI models and agents are probabilistic, so they sometimes hallucinate and reason poorly, generating mistakes. That means that adding generative AI to these products can be dangerous — merely replacing one risk with another.

## AllmaGen's Hybrid Approach To AI-Powered Decision-Making Overcomes The Challenge

Instead of relying on AI models or agents directly, we subsume these probabilistic technologies into a proprietary (patent-pending) deterministic framework we've developed. This approach makes the most of their powerful capabilities while avoiding their shortcomings. It's model-agnostic, so it can switch between several foundation models (such as OpenAI's Anthropic's models), reducing their tendencies to hallucinate and reason poorly when faced with high complexity.

The framework is based on what we call *strategy templates,* which we hone for specific domains. Strategy templates decompose complex tasks into steps while constraining deep reasoning and memory so that decomposition is reliable. And they model, quantify, and minimize risk, generating audit-ready decision justification documents.

Building on this safer framework, we've created three reliable domain-specific solutions based on strategy templates so far: Project Manager, RFP Manager, and Research Author — with more on the way.

As a result, AllmaGen lets companies use the power of AI to supercharge the business by making far better high-stakes decisions while sidestepping AI's risks — at a much lower cost than through fine-tuning foundational models or building custom ones.

## High-Stakes Decision-Making Is The Most Valuable — And Dangerous — Use of AI

AI's greatest promise is in boosting the speed and accuracy of high-stakes, high-risk decision-making in complex domains that range from supply chain logistics to RFP and RFQ management, technology procurement, and project management — each with millions and sometimes billions of dollars on the line. But with that promise comes peril.

- **It doesn't matter how efficiently AI helps you operate if you're steering the business in the wrong direction.** The widespread fascination with using AI to automate tasks is understandable since it cuts costs, improving the bottom line. But making higher-quality major decisions has far more business impact. That impact is not just on ensuring efficiencies make a meaningful difference but especially on the company's strategic direction and top line — boosting revenues and mitigating risks. And it applies not only at the inception of business initiatives but on an ongoing basis, for monitoring performance and course-correcting in response to change.

- **AI investments should prioritize decision quality over execution efficiency.** If a high-stakes decision is right, it can supercharge revenues and propel decision-makers' careers. If it's off the mark, it can undermine profits and ruin careers. That's why smart companies are focusing their AI investments on decision intelligence: improving decisions through data, modeling, and analytics — now also underpinned by AI.

- **The potential ROI is highest in complex, high-stakes industries.** Consider construction and engineering titans like Bechtel or Skanska that take on colossal projects like building airports, hospital complexes, refineries, railways, and sports arenas — whose cost can range from hundreds of millions to billions of dollars. They receive dozens of RFPs (requests for proposals) every day. And they face high-stakes decisions about each one — starting with whether to bid at all and, if they do, what to communicate in their RFP response, how, and at what price. And consider companies producing documents to *support* high-risk decisions, like clinical study reports from AstraZeneca, equity research from JPMorganChase, and IT vendor selection reports from Forrester Research. If they steer decision-makers in the wrong direction, the consequences can be catastrophic.

- **But adding AI models and agents to decision-making products can merely replace one risk with another.** Many decision-making technology providers are injecting AI into their platforms and tools — LLMs (large language models), LRMs (large *reasoning* models) and AI agents. But because these technologies are probabilistic, the risks they eliminate are replaced by different risks: hallucination and faulty reasoning. And agent-based approaches' repeated AI model invocations can lead to a cascading sequence of errors that compound — causing risk to balloon rather than shrink.

## AllmaGen's Hybrid Approach To AI Overcomes The Challenge

It's well established that an LLM's inaccuracies and inconsistencies tend to increase when incoming prompts are broad, since lower specificity causes it to improvise — probabilistically "filling in the blanks." As the reasoning that LRMs perform, recent research has shown that they can exhibit "reasoning collapse" when complexity is high. As a result, agent-driven systems based on these models exhibit the same flaws. These errors may be tolerable for some decisions, where the stakes are low and mitigation is therefore not worth the effort, but not when they're high.

AllmaGen has developed a hybrid approach (patent-pending) to DI that has proven exceptionally effective at sidestepping these problems. Rather than relying on AI models or agents directly, we subsume these probabilistic technologies into a proprietary deterministic framework we've developed that uses graph-of-thought techniques and narrowly scoped retrieval-augmented generation (micro-RAG). This approach makes the most of their powerful capabilities while avoiding their shortcomings. Specifically, we:

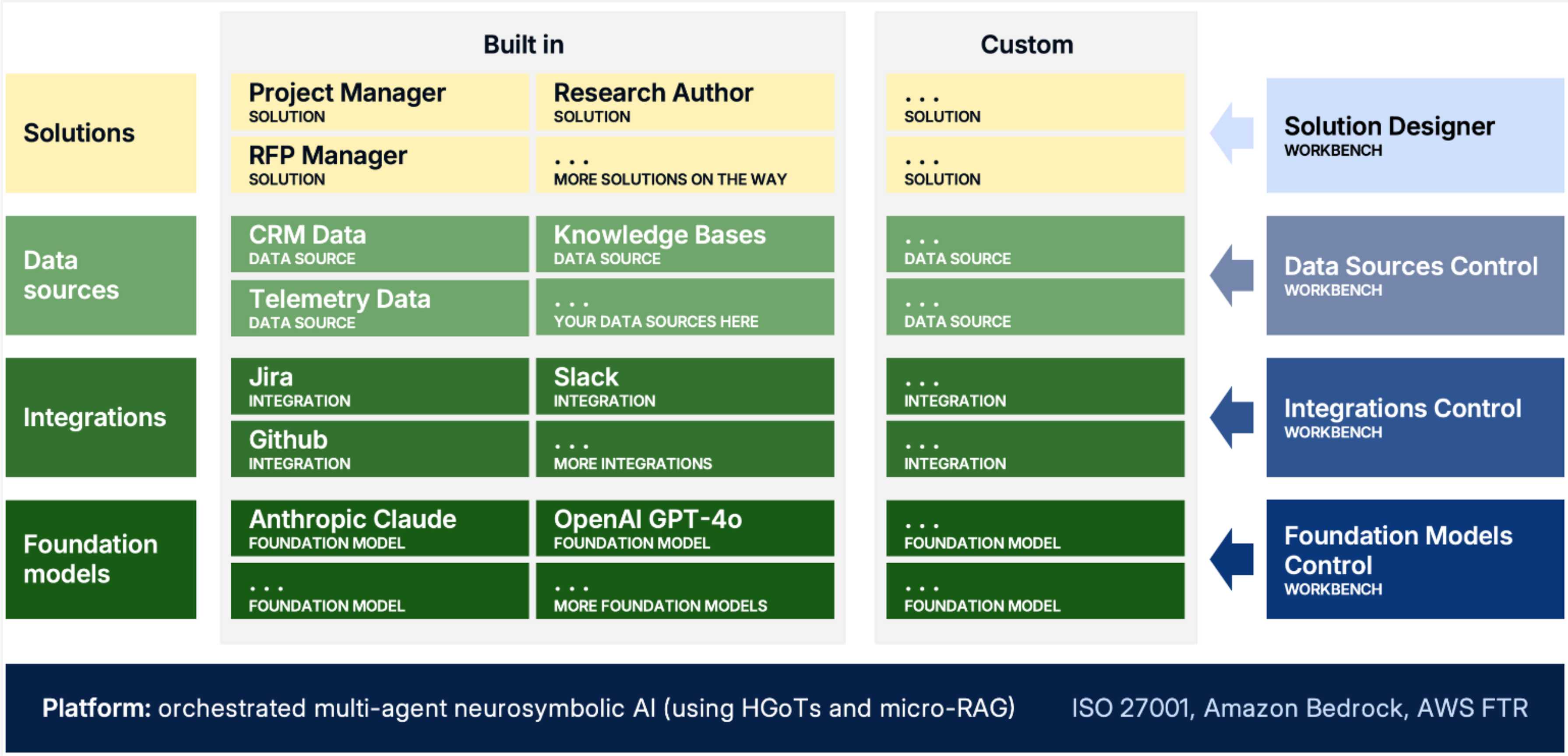
1. Design domain-specific strategy templates.
2. Decompose complex tasks into steps.
3. Constrain deep reasoning and memory.
4. Model, quantify, and minimize risk.
5. Generate audit-ready decision justification documents.

## 1. Design Domain-Specific Strategy Templates

Each strategy template we build is tailored to the unique requirements of a specific use case in a specific industry or function. It lets the system perform deep reasoning that invokes deterministic strategies, which in turn drive agents in a more controlled manner, so that the risks inherent in probabilistic AI techniques are contained and minimized.

We bundle each set of strategy templates we design for a specific industry or function into a reliable solution — such as our RFP Manager solution, Project Manager solution, and Research Author solution. And our platform enables rapid development of new strategy templates and solutions (see Figure 1).

**Figure 1:** AllmaGen includes solutions for specific industries and functions, built on our underlying platform



To develop a new strategy template as part of an industry-specific solution, we:

- **Research  the needs of professionals in a specific industry and function.** We work with professionals whose requirements the solution is intended to address, collaboratively researching and analyzing their needs.
- **Design each strategy template's steps and parameters based on our findings.** Based on the findings from our research and analysis, we design a deterministic sequence of steps, each of which takes unambiguously defined parameters as inputs (see Figure 2). Each step processes those inputs and generates appropriate outputs — using a mix of generative AI, machine learning, latent code generation, and other techniques, depending on the task. We chain these steps into linear, branching, and/or iterating flows as appropriate. In addition, the system provides lifecycle management of templates in production, including version control.
- **Present strategy templates in an organized catalog.** The system includes a catalog of hundreds of strategy templates out of the box (see Figure 3). And when it selects one of them in response to a prompt, it invites the user to review its selection before proceeding (see Figure 4). Once a user approves the system's selection of a strategy template, the system creates an instance and executes it.
- **Enable users to create new strategy templates.** The platform also includes a workbench for defining new strategy templates. This lets our customers adapt to new or modified enterprise processes in a fraction of the time that would be required if performing custom development.

**Benefit:** The resulting solutions and strategy templates reflect the real-world business requirements and expectations of domain professionals, and deliver trustworthy, reliably high accuracy.

**Figure 2:** One of the strategy templates for RFP management: RFP Bid Decision (screenshot)

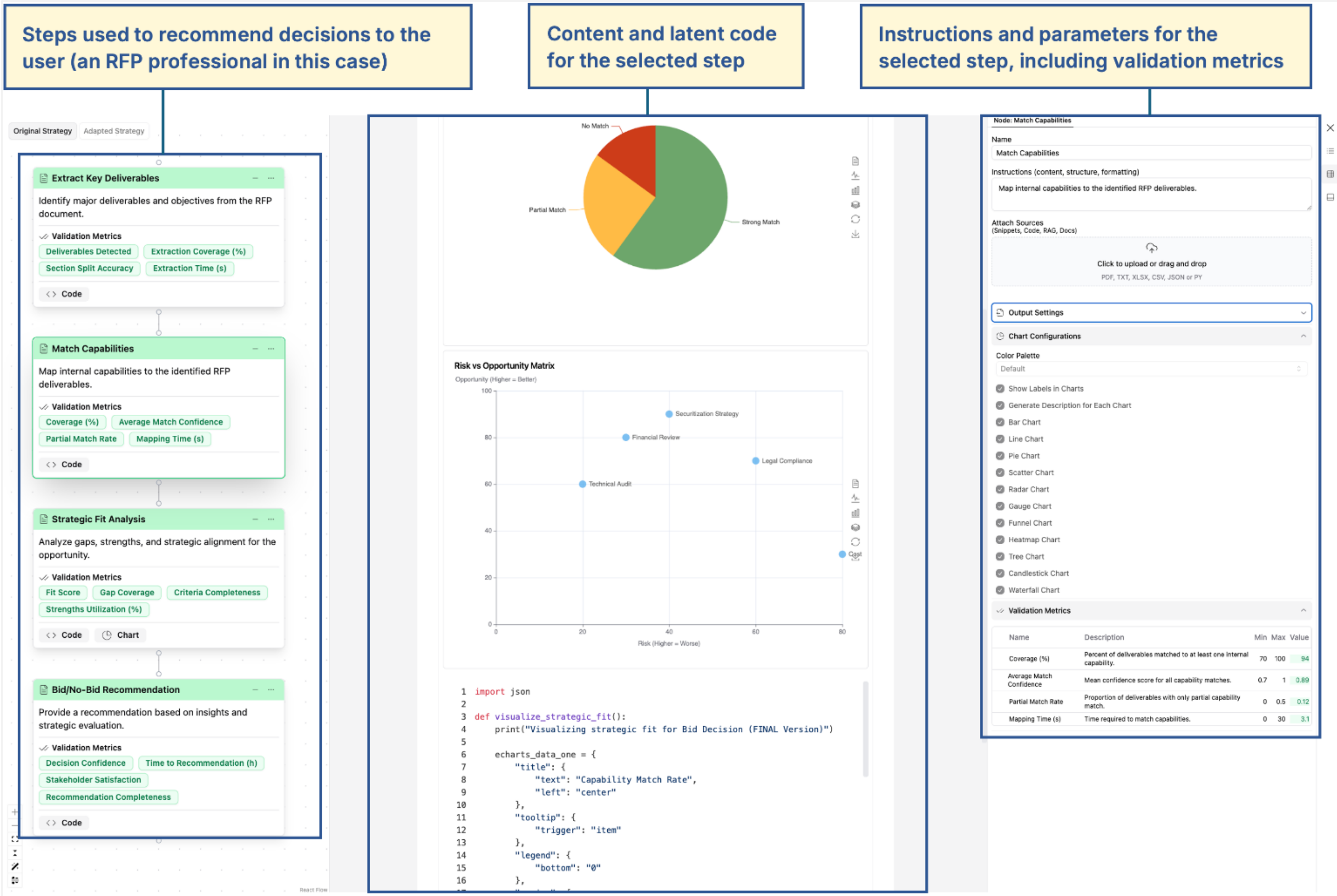


**Figure 3:** The strategy template catalog (screenshot)

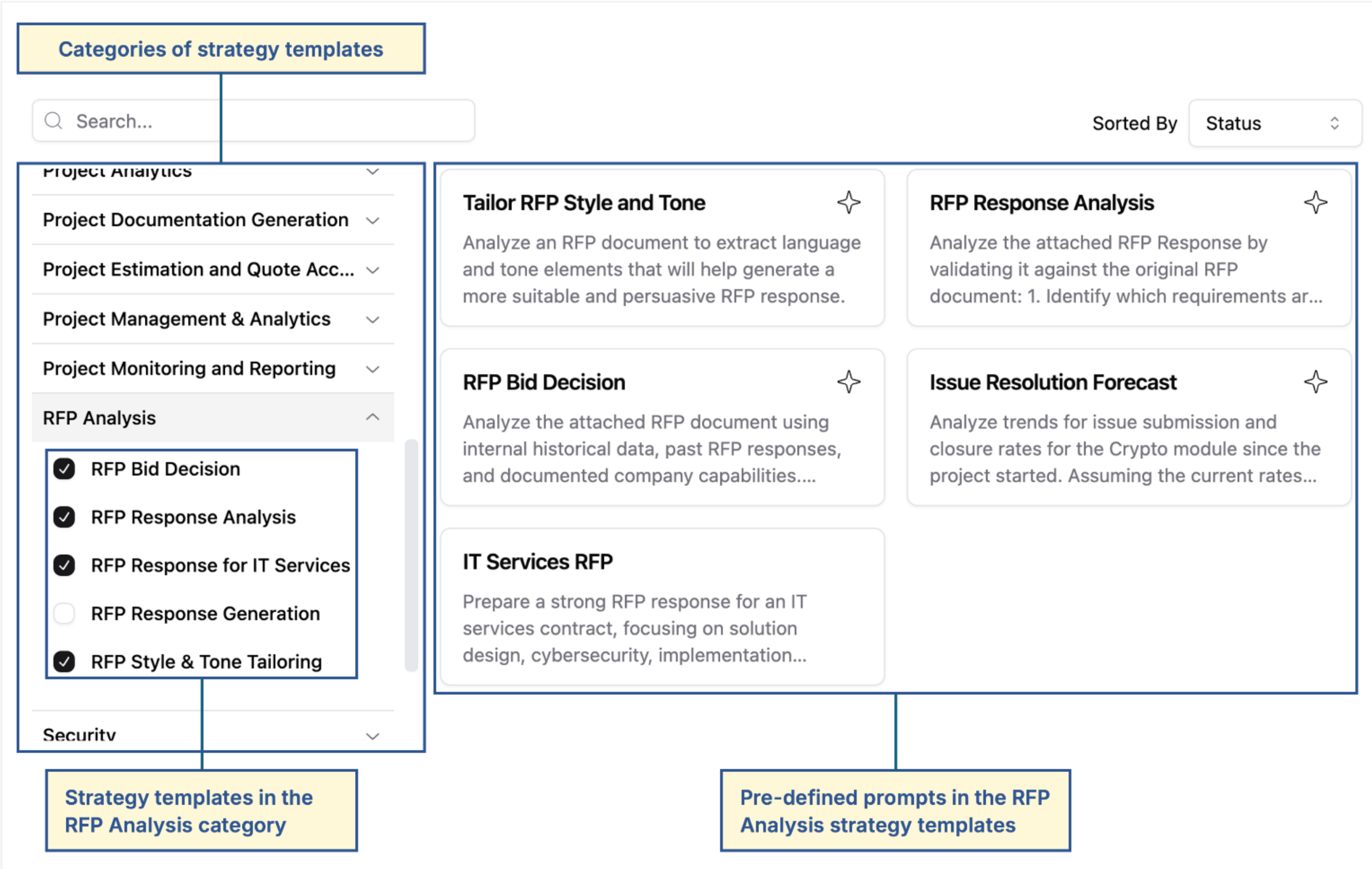
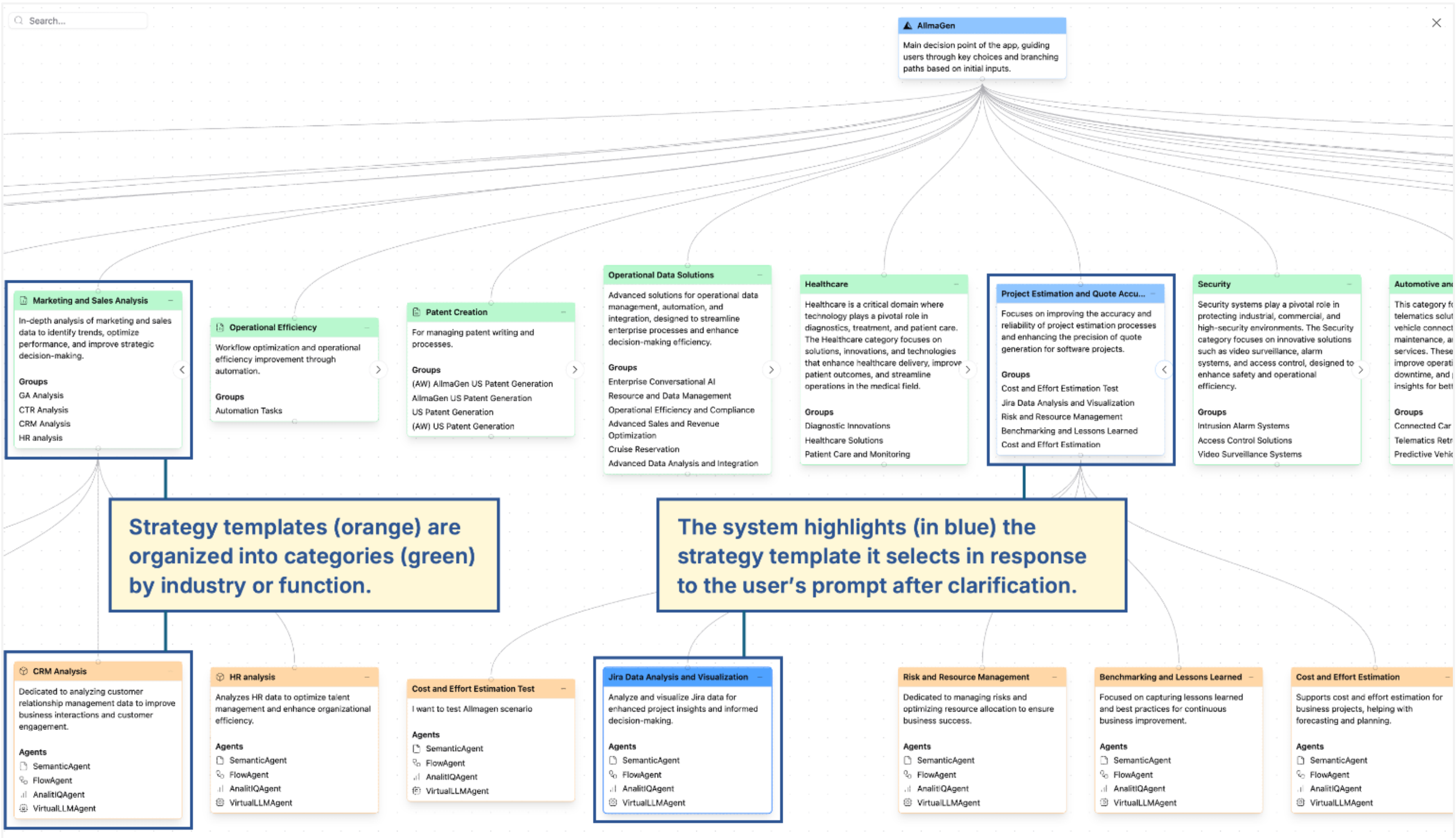


**Figure 4:** The strategy template runtime review and selection screen (screenshot)



## 2. Decompose Complex Tasks Into Steps

Instead of submitting a user's prompt to an LLM or LRM directly, AllmaGen starts by asking the user questions, to clarify their intent. Then, if the prompt is complex, the system breaks it down into simpler subtasks. However, it does so in a unique way that ensures the breakdown is correct and results in accurate and consistent results..

- **Avoid AI models' unreliability for direct task decomposition.** For high-stakes tasks, companies' risk tolerance is low. Therefore the system does not merely ask the model to break down a task without providing guidance — an approach that can cause it to invent subtasks that don't map to real needs.
- **Instead, steer the model via deterministic strategy templates.** Upon receiving a user prompt, the system examines its catalog of strategy templates and identifies the closest match. If the match exceeds a certain confidence threshold, the system proposes it to the user. If the user then approves it, the system's orchestrator then executes it. Since it contains a pre-validated sequence of templated steps, execution is predictable and reliable.
- **Alternatively, if no strategy template fits the prompt, steer the model based on examples.** If the match does not exceed a certain confidence threshold, the platform submits the prompt to the model along with examples that are aligned with the user's intent. The examples show effective mappings from complex prompts to subtasks. In addition to these two items, it includes instructions not to execute the prompt but to break it down into simpler prompts similarly to the breakdowns illustrated by the examples. AllmaGen then submits the simpler prompts to the model individually via specialized agents, along with instructions about limiting scope.

**Benefit:** AI models are far less likely to make errors when prompts describe simpler tasks, since there is less need to "fill in the blanks" — this is a well-understood benefit of task decomposition. Additionally, however, AllmaGen's strategy templates 1) cause the model's task decomposition process to snap to the approaches suggested by the examples and 2) isolate and tailor risk and accuracy assessments via separate specialized agents applied to these simpler sub-tasks.

## 3. Constrain Reasoning and Memory

What is called "reasoning" in a generative AI context is not what is meant by "reasoning" in common parlance. Specifically, generative AI models do not perform logical operations on premises to derive conclusions — they generate pseudo-reasoning: probabilistic token sequence completions that bear a statistical resemblance to the language of logical reasoning. So whereas reasoning in the normal sense cannot, by definition, draw false conclusions from true premises, an LLM's pseudo-reasoning can go off the rails while sounding logical. (See the recent research [1] about this mentioned above.)

In addition, LLMs suffer from a well-known propensity to veer into hallucination due to two factors: distraction and semantic drift. These occur when the context window includes material that is extraneous — not aligned with the purpose of the intended subtask.

To overcome this problem, for each subtask-specific LLM invocation, we:

- **Exclude extraneous material from the global portion of the context window.** We ensure the global portion includes only the bare minimum that is relevant to all the subtasks, with no "crosstalk."
- **Inject only subtask-relevant material in the local portion of the context window.** We construct this portion parsimoniously, including semantically relevant inputs and no others.

**Benefit:** These measures significantly improve inference and semantic precision, and promote coherent, task-consistent outputs in multi-step reasoning.

## 4. Model, Quantify, and Minimize Risk

We model, quantify, and minimize decision risk at two stages:

- While designing strategy templates.
- At runtime, when the system applies a template in response to a specific user prompt.

Much of the risk management logic is specific to individual strategy templates or even to particular subtasks — for example to verify compliance with legal requirements or not contradict known enterprise operating rules.

## Validation of Strategy Templates During Design

During the process of developing a strategy template, we:

- Validate it through iterative simulation over domain-specific input sets. This is an additional layer above and beyond model training or fine-tuning: it examines the LLM's response to estimate and analyze risks.
- Perform Bayesian modeling and Monte Carlo sampling to estimate probability of success across varying conditions.
- Stress-test failure modes, transition blockages, and edge case branches for resilience.

## Runtime Evaluation of Strategies and Decisions

In each strategy template, we embed multiple deterministic testing and auditing techniques that mitigate risk, for the strategy overall and for each step. Every time the system executes a strategy in production, it scores each step against:

- Dynamic context relevance.
- Resource availability and latency tolerance.
- Stepwise inference confidence.
- Historical pathway success likelihood.

The system applies a unified utility function to detect excessive risks, and proceeds with path expansion conditionally: if the function identifies excessive risk, it either re-routes to predefined mitigation steps or halts execution of the strategy and notifies the user about the reason why.

In addition, the system includes "human-in-the-loop" mechanisms so that decision-makers can examine interim results and either validate or course-correct. The system also preserves a detailed audit trail to ensure its reasoning and recommendations are explainable in a form suitable for auditing.

## Compute Evaluation Metrics

To minimize risk, AllmaGen steers execution based on several key metrics it calculates, including:

- Variations between the template and the generated graph in terms of number of nodes, stages, properties, etc.
- Variations between the prompt and the generated graph in terms of keyword coverage and length, as well as match percentage.
- Semantic similarities between prompts and node descriptions and between templates and generated graph descriptions.

**Benefit:** The uncertainty and risk inherent in a decision are both well understood and minimized, prior to acting on a recommendation. But the process takes mere minutes, instead of the hours, days, or weeks it would require if analyzed manually. This gives decision-makers access to more opportunities while minimizing exposure to risk.

## 5. Generate Audit-Ready Decision Justification Documents

When AllmaGen provides answers and makes recommendations, it always also explains the rationale for its outputs and it records that reasoning, the validation metrics it applies (see Figure 5) and its recommendations in static, persistent assets, including documents, charts, and dashboards (see Figure 6), including risk analysis at the strategy and subtask level.

**Benefit:** Enterprises can examine the system's recommendations and the reasoning that underlies them at any time, including long after the task has been performed. This ensures explainability and auditability, as an enterprise uses AllmaGen to make its high-stakes decision-making faster and more accurate.

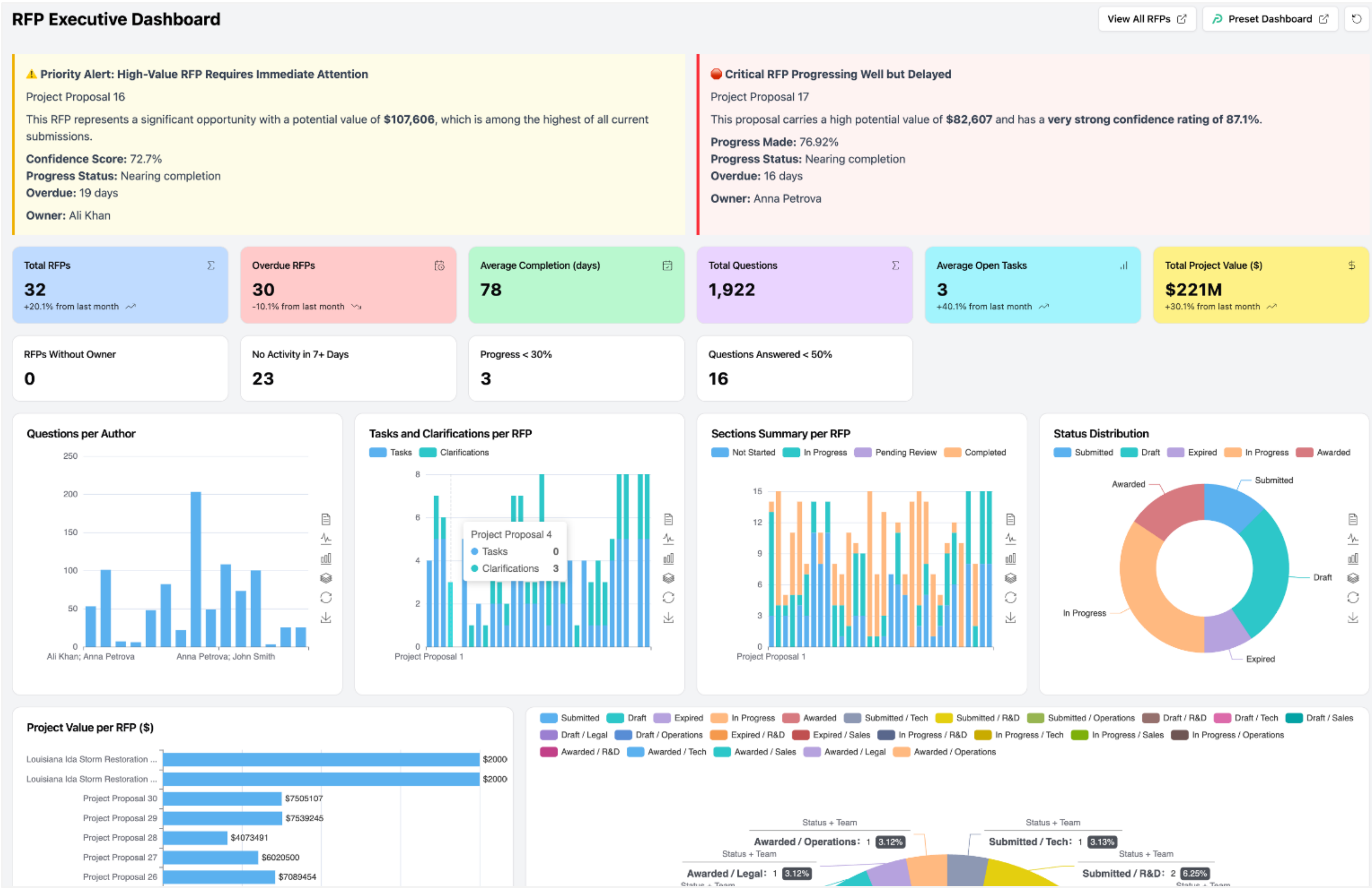
**Figure 5:** AllmaGen shows the validation metrics it applies (screenshot)

## RFP Validation Metrics

🔍 **Risk Scoring Factors:**

| Step | Metric Name | Description | Min | Max | Value |
|---|---|---|---|---|---|
| Extract Key Deliverables | Deliverables Detected | Number of unique deliverables detected in RFP. | 3 | 20 | 5 |
| | Extraction Coverage (%) | Percent of RFP sections covered by at least one deliverable. | 60.0 | 100.0 | 82.0 |
| | Section Split Accuracy | Correctness of auto-section splitting. | 0.7 | 1.0 | 0.9 |
| | Extraction Time (s) | Time taken for deliverable extraction. | 0.0 | 20.0 | 6.0 |
| Match Capabilities | Coverage (%) | Percent of deliverables matched to at least one internal capability. | 70.0 | 100.0 | 94.0 |
| | Average Match Confidence | Mean confidence score for all capability matches. | 0.7 | 1.0 | 0.89 |
| | Partial Match Rate | Proportion of deliverables with only partial capability match. | 0.0 | 0.5 | 0.12 |
| | Mapping Time (s) | Time required to match capabilities. | 0.0 | 30.0 | 3.1 |
| Strategic Fit Analysis | Fit Score | Composite strategic fit score. | 0.7 | 1.0 | 0.81 |
| | Gap Coverage | Proportion of gaps identified relative to potential gaps. | 0.75 | 1.0 | 0.93 |
| | Criteria Completeness | Percent of fit criteria covered. | 85.0 | 100.0 | 96.0 |
| | Strengths Utilization (%) | Percent of strengths actually used in fit analysis. | 60.0 | 100.0 | 77.0 |
| Bid/No-Bid Recommendation | Decision Confidence | Estimated confidence in the bid/no-bid recommendation. | 0.7 | 1.0 | 0.79 |
| | Time to Recommendation (h) | Elapsed time to produce a recommendation. | 0.0 | 24.0 | 5.2 |
| | Stakeholder Satisfaction | Stakeholder satisfaction with the process. | 0.8 | 1.0 | 0.91 |
| | Recommendation Completeness | Percent of recommended analysis steps completed. | 90.0 | 100.0 | 98.0 |

**Figure 6:** An example AllmaGen executive dashboard (screenshot)



## AllmaGen Brings High Reliability To High-Stakes AI-Powered Decision-Making

AllmaGen guides decision-making by connecting and analyzing all available data including historical precedents, evaluating the risk of a decision from multiple angles, and presenting its findings and recommendations via interactive executive dashboards as well as detailed documents and charts. And once a high-stakes decision is made, the process is not over. Decision-makers must then reassess and course-correct over time — estimating whether projects risk running late, are inadequately resourced, and so on. AllmaGen helps continuously monitor these risks as well, so that decision-makers can make adjustments as needed.

By using AllmaGen for decision-making (in addition to automation), companies can make high-stakes decisions with confidence. It's what the three solutions we've developed so far are designed for: RFP Manager, Project Manager, and Research Analyst. And the platform enables building new strategy templates and solutions rapidly. That means that in addition to automating business processes and workflows reliably, you can seize the upside of promising business opportunities rather than leaving money on the table, and avoid the costly downside of dangerous mistakes, steering the business and its efficiency gains in the right direction.

---

**Interested in learning more about AllmaGen? Visit our website ([www.allmagen.com](www.allmagen.com)) and contact us for a discussion and demonstration.**

---

This whitepaper was authored by David Truog ([PragmaticMind.ai](PragmaticMind.ai)).

---