# ARCEE ORCHESTRA

## Purpose-built AI agents. Powered by industry leading SLMs.

Arcee Orchestra makes it easy to build custom AI workflows. Automatically route tasks to specialized small language models (SLMs) and create AI agents for complex tasks.

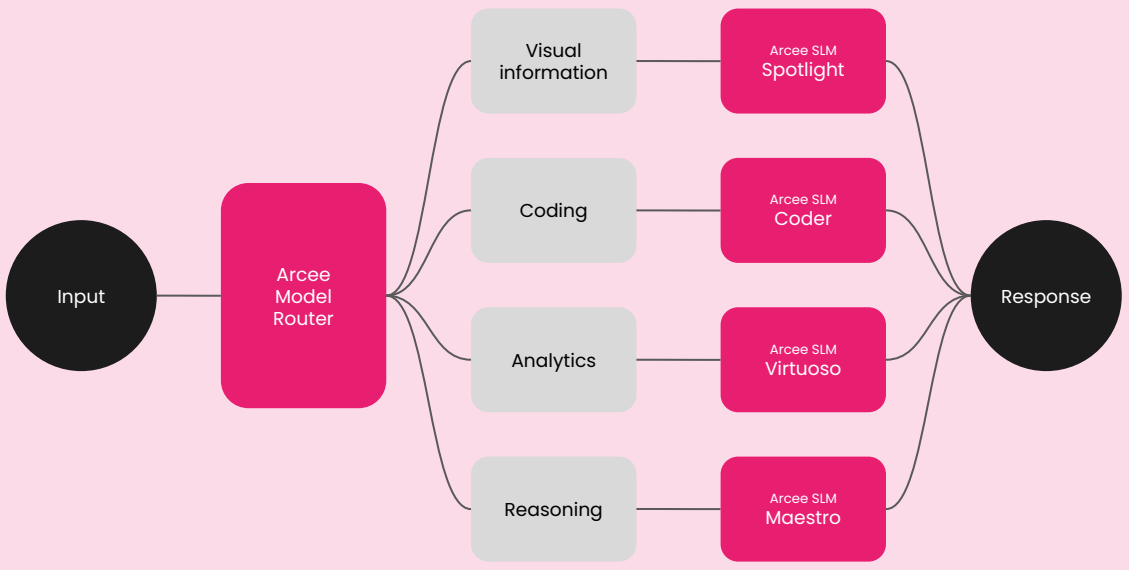
### When SLMs work together, you get better results

Arcee Orchestra breaks queries into tasks, routes each one to the right SLM, then consolidates a response.

You get detailed, trustworthy answers—fast.

### Arcee Orchestra is Enterprise AI built to work

Use specialized SLMs to handle multi-step tasks, eliminate busywork, and get better outputs, faster.




## Why Arcee Orchestra?

### Task-focused AI without the complexity

Arcee AI SLMs understand your data, workflows, and goals—so you get reliable, powerful tools with none of the bloat of an LLM.

### Scalable, manageable AI

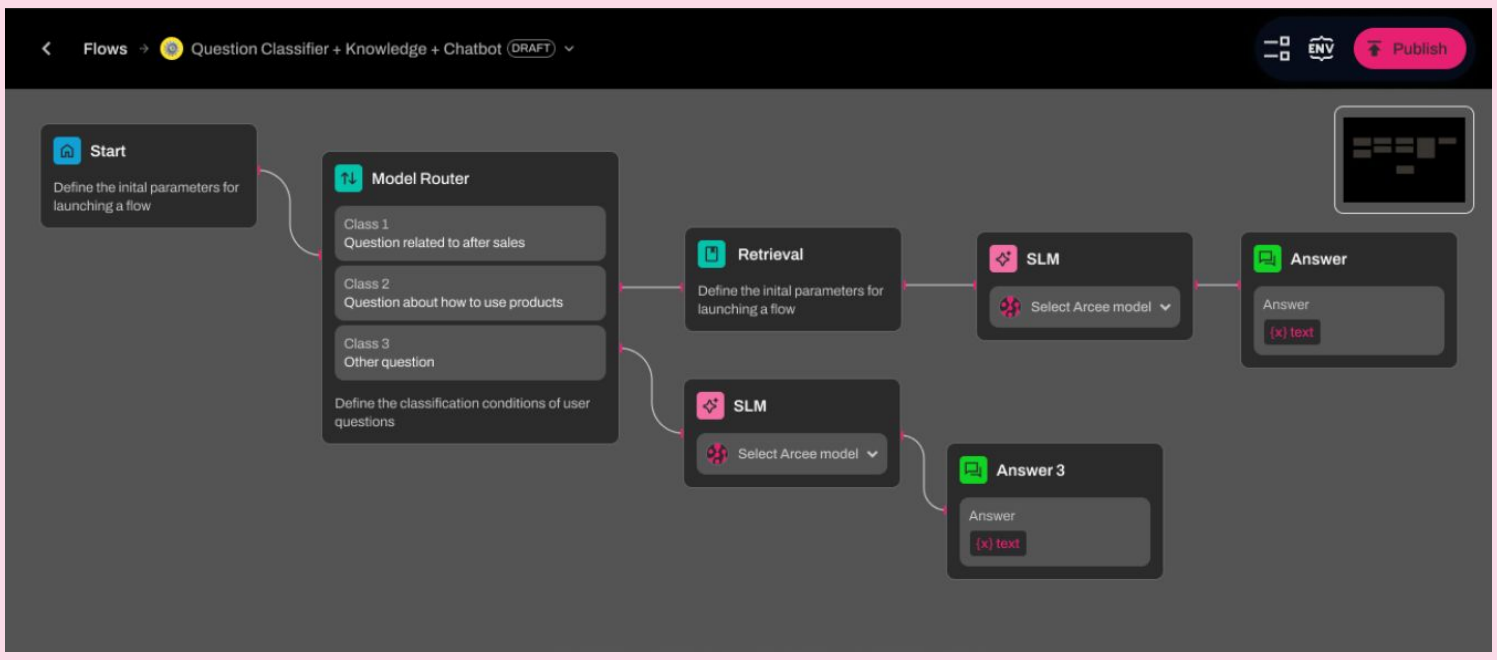
Automate repetitive tasks and empower teams while maintaining full control over AI actions and decisions.

### An AI model to match every job

Intelligent routing assigns each task to the SLM best suited for it. You get accurate, reliable, trustworthy AI.

### Deploy in any environment

Run Arcee Orchestra anywhere: On-prem, in your VPC, or in our cloud. Maintain the right level of data security for your organization.




### Create an AI workflow in clicks

Our no-code interface allows anyone to create sophisticated custom workflows.

Start with our templates, customize based on your processes and data, and get useful tools that radically speed up productivity.

Enjoy flexible deployment and easy integration, with full control of your data.

### Arcee Orchestra Integrates with

asana · Jira · Discord · Slack · GitHub · Dropbox · Google Calendar · Google Sheets · salesforce · Shopify · LinkedIn · Calendly

Ready to learn more about how a symphony of Small Language Models can transform your business? Book an Arcee Orchestra demo today!

# MODEL ENGINE

arcee.ai

## The small language models (SLMs) that power Arcee Orchestra.

| MODEL NAME | WHAT IT EXCELS AT | DESCRIPTION | TOP USE CASES |
|---|---|---|---|
| Virtuoso Large | General Purpose | Our most powerful and versatile general-purpose model, designed to excel at handling complex and varied tasks across domains. With state-of-the-art performance, it offers unparalleled capability for nuanced understanding, contextual adaptability, and high accuracy. | • Advanced content creation, such as technical writing and creative storytelling<br>• Data summarization and report generation for cross-functional domains<br>• Detailed knowledge synthesis and deep-dive insights from diverse datasets<br>• Multilingual support for international operations and communications |
| Virtuoso Medium | General Purpose | A versatile and powerful model, capable of handling complex and varied tasks with precision and adaptability across multiple domains. Ideal for dynamic use cases requiring significant computational power. | • Content generation<br>• Knowledge retrieval<br>• Advanced language understanding<br>• Comprehensive data interpretation |
| Virtuoso Small | General Purpose | A streamlined version of Virtuoso, maintaining robust capabilities for handling complex tasks across domains while offering enhanced cost-efficiency and quicker response times. | • General-purpose task handling<br>• Business communication<br>• Automated document processing for mid-scale applications |
| Caller Large | Tool & Function Calls | Engineered for seamless integrations, Caller-Large is a robust model optimized for managing complex tool-based interactions and API function calls. Its strength lies in precise execution, intelligent orchestration, and effective communication between systems, making it indispensable for sophisticated automation pipelines. | • Managing integrations between CRMs, ERPs, and other enterprise systems<br>• Running multi-step workflows with intelligent condition handling<br>• Orchestrating external tool interactions like calendar scheduling, email parsing, or data extraction<br>• Real-time monitoring and diagnostics in IoT or SaaS environments |
| Coder Large | Coding | A high-performance model tailored for intricate programming tasks, Coder-Large thrives in software development environments. With its focus on efficiency, reliability, and adaptability, it supports developers in crafting, debugging, and refining code for complex systems. | • Writing modular, reusable code across various programming languages<br>• Debugging and optimizing performance in large-scale applications<br>• Generating efficient algorithms for computationally intensive tasks<br>• Supporting DevOps processes, such as script automation and CI/CD pipelines |
| Coder Small | Coding | A compact, high-performance coding model designed for efficient programming tasks, including generating code, debugging, and optimizing scripts for smaller projects. | • Lightweight development tasks<br>• Automated code reviews<br>• Generating templates or prototypes quickly, code completion |
| Spotlight | Vision-Language | Spotlight is a specialized vision-language model designed to bridge the gap between visual and textual data. It excels at interpreting images, generating insights, and combining multimodal data to make complex information more accessible. | • Analyzing visual data for industry-specific applications like medical imaging, architecture, or fashion<br>• Multimodal storytelling through synchronized text and visuals<br>• Automated annotation of large-scale image datasets for AI training<br>• Extracting insights from charts, graphs, and infographics for business analytics |
| Maestro | Reasoning Engine | Maestro is the pinnacle of reasoning and analytical AI, designed to tackle the most complex problem-solving scenarios with unparalleled precision and depth. It excels in handling multifactorial decision-making, abstract reasoning, and scenario modeling, delivering insights that rival expert-level human analysis. | • Strategic decision-making for high-stakes scenarios, such as financial modeling or geopolitical analysis<br>• Comprehensive diagnostics and root-cause analysis across technical and non-technical domains<br>• Advanced simulation modeling for predicting outcomes or evaluating "what-if" scenarios<br>• Thought leadership assistance for crafting well-reasoned arguments, strategies, and long-term plans |

## And Featuring

| | | | |
|---|---|---|---|
| Metronome | Intelligent Router | A small yet highly efficient 300M parameter model trained to route messages to the most appropriate model in the Arcee Orchestra lineup, based on the context of the input. | • Query classification<br>• Task delegation<br>• Operational efficiency in hybrid workflows involving multiple models |