

PERSONAL AI THAT STAYS YOURS

# Your AI. Your Data. Your Models.

Run local AI, connect providers you already use,  
and keep your work, history, and model choice  
together.

---

One workspace where approved context helps your models and  
Hermes Bots understand how you work.

**ONE CLICK** LOCAL AI  
**ONE PLACE** FOR EVERY MODEL  
**ONE CLICK** HERMES BOTS IN PLUS

wonderwavetech.com

**+852 9880 5017**

wonderwavetechnologies@proton.me

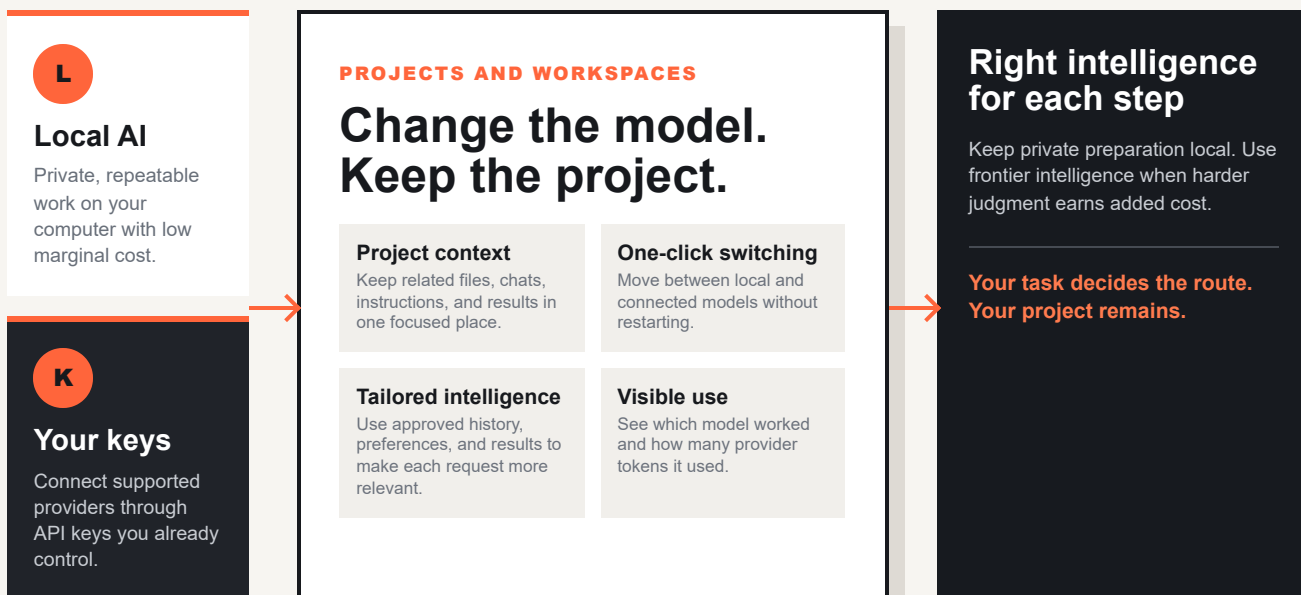


**ONE WORKSPACE, ANY MODEL**

# Your models. **One place to use them.**

Organize work into projects and workspaces. Install local AI, connect providers you already use, and change models while each project keeps its files, conversations, instructions, and useful context together.

**What changes:** AI work gains a durable home. Projects keep related files, conversations, instructions, decisions, and results together. Approved context can strengthen each new chat or Bot without surrendering your history to one provider.



## Optimized for your machine

Velociti profiles available CPU memory, GPU memory, and acceleration, then tunes model fit, quantization, context, cache, memory placement, decoding, and batch settings.

## Local History you own

Search and reopen prior conversations without depending on a provider-owned chat account. Useful work remains available when models change.

## Token count shown

Per-message receipts show model, provider, input, output, cache, and reasoning tokens so usage is visible while you work.

## Built-in CDP Listener

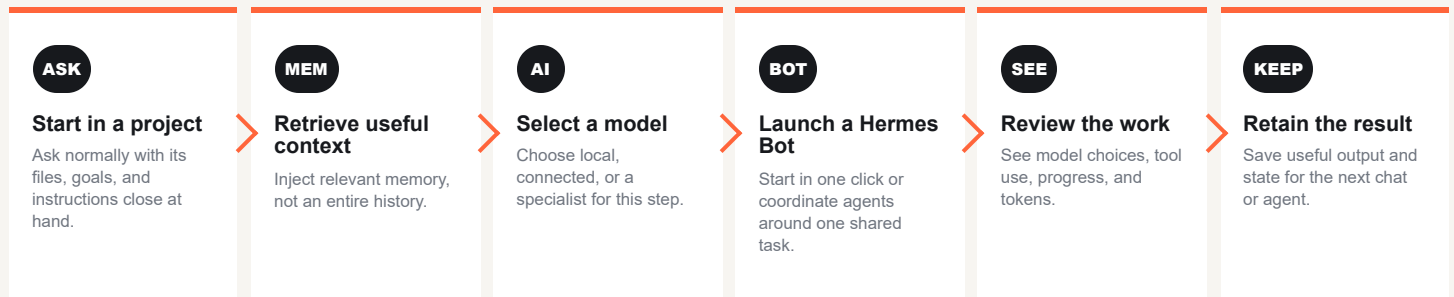
A local Chrome DevTools Protocol channel lets Velociti inspect visible IDE actions and verify work without relying on screen scraping.

**How it works:** Velociti installs a compatible local runtime, measures your hardware, selects safe model and memory settings, validates performance, and adds the tuned model to the same picker as your connected providers.

**WORK THAT CONTINUES**

# Chats become work **that can continue.**

Work directly in chat or create a Hermes Bot in Plus. Velociti carries selected context, model choice, agent state, token use, and accepted results from one step to the next.



## Projects with useful memory

Each project keeps its files, conversations, instructions, decisions, preferences, corrections, and accepted results together. Velociti selectively brings useful context into future chats and Bots without changing the base model or exposing your full history.

## Shared workspaces in Plus

Invite collaborators into a common project where shared files, conversations, context, Bots, and accepted results remain connected. People and agents can continue the same work without rebuilding its background.

## Hermes Bots in Plus

Create a bot in one click for research, summaries, monitoring, or other repeatable work. Choose its models, tools, files, memory, and schedule. See progress and token use, pause it, or continue from saved state.

**SPEND INTELLIGENCE WHERE IT MATTERS**

## Local handles routine. Frontier handles exceptions.

A frontier model can supervise or review work prepared by a local model. Hard judgment receives stronger intelligence. Private and repeatable steps stay local when appropriate.

**Token optimization reduces avoidable context.**

Selective memory injects only relevant history.

Tool definitions enter only when needed.

Irrelevant tools are pruned from each turn.

Live token counts show input, output, cache, and reasoning use.

**Result: models and Bots gain useful context from accepted work. Responses become more relevant, repeated prompting falls, and your tailored intelligence survives model changes.**

**CHOOSE HOW YOU WORK**

# Start free. Add continuity when AI becomes part of your work.

Free gives individuals projects and a personal AI workspace. Plus adds shared workspaces, collaboration, persistent automation, orchestration, and deeper visibility for recurring work.

Both products protect model choice and keep useful work portable. Choose Plus when tasks need to continue beyond a conversation.

**DOWNLOAD AND SELF-INSTALL**

## Velociti Free

One familiar place to use personal AI.

- ✓ **Local AI tuned for your machine**  
Install in one click. Velociti profiles your hardware, tunes model fit and runtime settings, validates performance, and adds it to your picker.
- ✓ **Your models and your keys**  
Use local models and connect supported providers through BYOK.
- ✓ **One-click model switching**  
Change intelligence without rebuilding the task.
- ✓ **Personal workspaces and projects**  
Keep each body of work focused with its own files, chats, instructions, context, and results.
- ✓ **Searchable Local History**  
Reopen prior conversations without relying on a provider-owned chat history.
- ✓ **Token counts and optimization**  
See token receipts while selective memory and tool pruning reduce avoidable context.
- ✓ **Built-in CDP Listener**  
Inspect visible IDE actions and verify work through a local protocol channel.
- ✓ **User-approved tool access**  
Choose what an agent may read, change, or call.

**FOR ONGOING AND COORDINATED WORK**

## Velociti Plus

Create and monitor Hermes Bots that keep useful work moving beyond the chat.

- ✚ **Shared workspaces and projects**  
Bring people into a common workspace with shared files, project context, conversations, agents, and accepted results. Everything in Free remains included.
- ✚ **One-click Hermes Bot setup**  
Give a bot a task, choose what it may use, and start without assembling a separate automation stack.
- ✚ **Hermes monitoring and schedules**  
Run work while you are away. See progress, model use, tokens, results, and anything needing attention.
- ✚ **Project continuity**  
Carry relevant state across chats, people, and agents while the work remains attached to its project.
- ✚ **Cross-agent orchestration**  
Coordinate research, drafting, review, and checking around one shared task.
- ✚ **Persistent tasks and recovery**  
Pause, stop, restart, and continue longer work from saved state.
- ✚ **Bots strengthened by your context**  
Approved preferences, corrections, decisions, and successful outputs make recurring work more tailored over time. Selective injection supplies useful context without sending the entire history.

**Velociti does not ask you to choose one AI forever.**

**It gives you one place to use the right intelligence for the work.**