ALIBABA CLOUD
INTELLIGENCE GROUP

Introducing Next-gen Agent Infra
Alibaba Cloud AgentBay’s Global Product Launch

Alibaba Cloud Workspace

0 / MARKET ANALYSIS AND KEY CONCEPTS

The Age of Agents Demands Innovative Cloud Infrastructure

The Evolution of Agent Capabilities

From "Workflow" to "Autonomous Planning"

Traditional AI applications rely on preset rules and fixed processes. They are like robotic arms on an assembly line, only performing predefined tasks. In contrast, modern AI Agents possess powerful reasoning, learning, and adaptation capabilities. They can autonomously understand complex scenarios and objectives, dynamically plan execution paths, and flexibly adjust strategies in response to change. They don't just execute tasks; they understand why and decide how.

From "Q&A" to "Action"

While traditional large language models (LLMs), such as ChatGPT, focus on text-based Q&A and content generation, modern general-purpose Agents, such as Manus, AutoGLM, and Operator, deliver end-to-end Plan-Execute-Deliver capabilities. They can operate applications, manage files, and automate office tasks just like a human. These abilities require a real, interactive operating environment, far beyond what simple API calls or local inference can provide.

Surging Enterprise Demand

Enterprise Agent Adoption

IDC predicts that by 2026, 50% of China's top 500 companies will adopt AI Agents for data analysis and business automation, and 40% will have achieved unified governance of their AI and data.

Massive Productivity Gains

In real-world applications, Agents like DeepResearch and Coding can reduce complex tasks that once took 8 hours down to just 5-10 minutes, delivering a huge boost in productivity.

The Ideal Runtime Environment for Agents

Secure Isolation and Compliance

- Secure sandboxing: Agents run in isolated cloud environments, mitigating risks to local systems and preventing issues like malicious code execution or data leakage.
- Permissions & compliance: Enterprise-grade Agents require strict control over data access and permissions. A cloud environment facilitates centralized management and auditing to meet compliance standards (e.g., GDPR, internal corporate controls).

Elasticity & Cross-OS Support

- Elastic compute: The cloud can dynamically allocate compute resources, supporting massive concurrency and auto-scaling during peak loads, which significantly reduces enterprise IT costs.
- Cross-OS capabilities: Long-horizon tasks often require interaction across different operating systems. AgentBay natively supports Windows, Linux, and Android environments.

Multi-Tool Orchestration & Integration

- Multi-tool operation: Agents can operate a wide array of applications in the cloud—including browsers, Office suites, IDEs, and databases—to complete complex automation tasks.
- Rich toolchain: The cloud environment can be pre-configured with a rich set of tools and APIs, allowing Agents to call them on-demand to improve task success rates and efficiency.

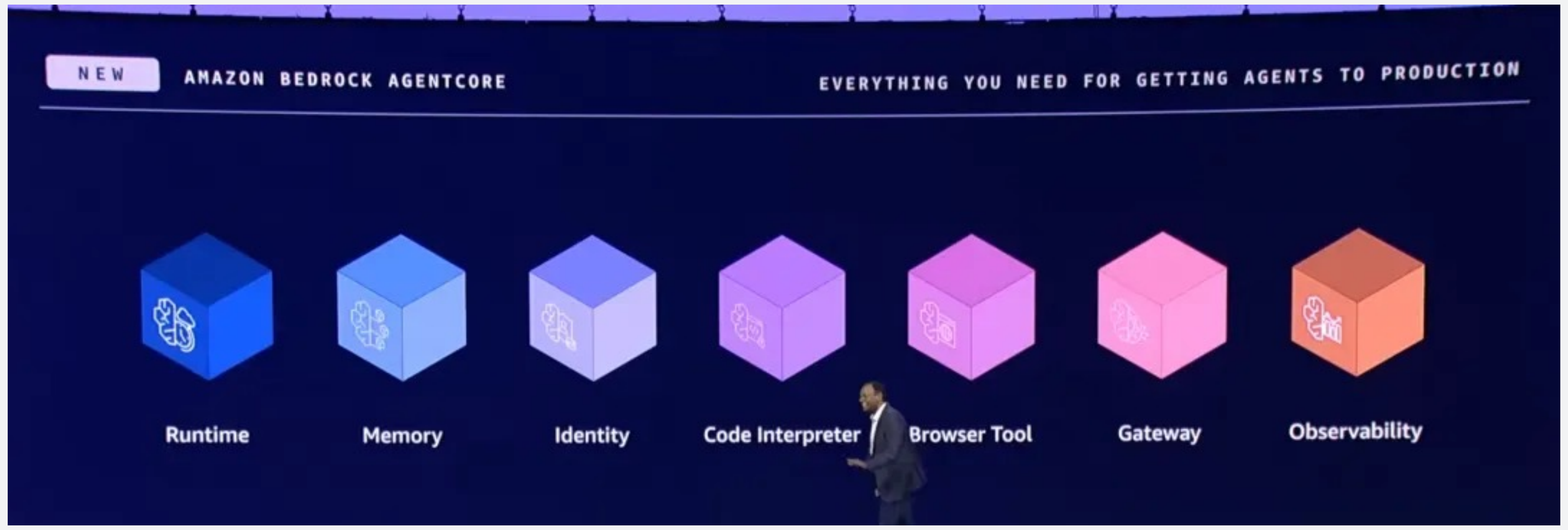
Observability & Automated O&M

- Task logging & traceability: The cloud environment facilitates detailed logging of Agent actions and task statuses, enabling automated capabilities like auto-retries on failure and anomaly alerting.
- Continuous integration & upgrades: The Agent platform can be centrally upgraded and maintained, allowing for rapid responses to new requirements and security vulnerabilities.

Market Heats Up: Major Players Enter as Unicorns Emerge

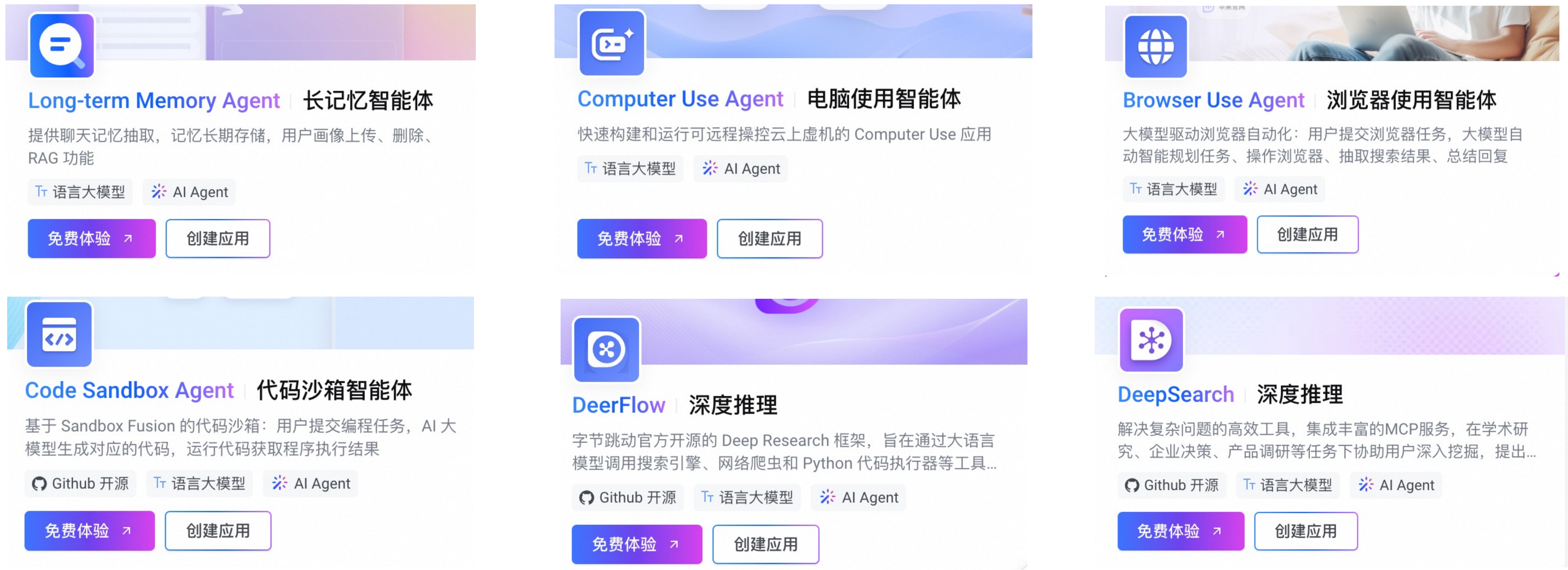
Sandboxes & subagents are becoming standard components, with significantly more players outside China than in China.

AWS AgentCore



Launched on July 17, 2025, AgentCore is a comprehensive agent development platform featuring seven major services, including identity, memory, and observability. Its sandbox and memory modules are direct competitors to AgentBay's current product roadmap.

Volcano Engine Cloud-Native Agent Suite


Long-term Memory Agent 长记忆智能体

提供聊天记忆抽取，记忆长期存储，用户画像上传、删除、RAG 功能

语言大模型 · AI Agent

免费体验 ↗ | 创建应用

Computer Use Agent 电脑使用智能体

快速构建和运行可远程操控云上虚机的 Computer Use 应用

语言大模型 · AI Agent

免费体验 ↗ | 创建应用

Browser Use Agent 浏览器使用智能体

大模型驱动浏览器自动化：用户提交浏览器任务，大模型自动智能规划任务、操作浏览器、抽取搜索结果、总结回复

语言大模型 · AI Agent

免费体验 ↗ | 创建应用

Code Sandbox Agent 代码沙箱智能体

基于 Sandbox Fusion 的代码沙箱：用户提交编程任务，AI 大模型生成对应的代码，运行代码获取程序执行结果

Github 开源 · 语言大模型 · AI Agent

免费体验 ↗ | 创建应用

DeerFlow 深度推理

字节跳动官方开源的 Deep Research 框架，旨在通过大语言模型调用搜索引擎、网络爬虫和 Python 代码执行器等工具...

Github 开源 · 语言大模型 · AI Agent

免费体验 ↗ | 创建应用

DeepSearch 深度推理

解决复杂问题的高效工具，集成丰富的MCP服务，在学术研究、企业决策、产品调研等任务下协助用户深入挖掘，提出...

Github 开源 · 语言大模型 · AI Agent

免费体验 ↗ | 创建应用

Released at the end of April 2025, this suite packages common sub-agents as standard services, covering automated sandboxes, memory, and deep reasoning.

Computer Use

OpenAI Operator: An integrated Computer Use service featuring a built-in sandbox and the GPT-4o vision model, designed to solve complex enterprise automation challenges. It is offered as a closed-source SaaS and API.

Anthropic: Provides a first-party tool, currently a Beta feature, available for Claude 3.5 models and later. It supports keyboard and mouse operations, system interface automation, and screen capture.

Browser Use

BrowserBase: A startup from late 2024 that developed the proprietary Stagehand browser automation framework. It offers an integrated Browser Agent solution and is valued at over $300 million.

HyperBrowser: A 2024 startup focused on headless browser automation in containerized environments. It provides an MCP tool and supports high concurrency, stealth mode, and automatic proxying.

Code Space

E2B: A code sandbox service popularized by Manus. It is open-source with a closed-source PaaS offering that has processed over 100 million calls in the last two months (hosted on AWS and GCP). It has raised $5.5 million in a seed round with an undisclosed valuation.

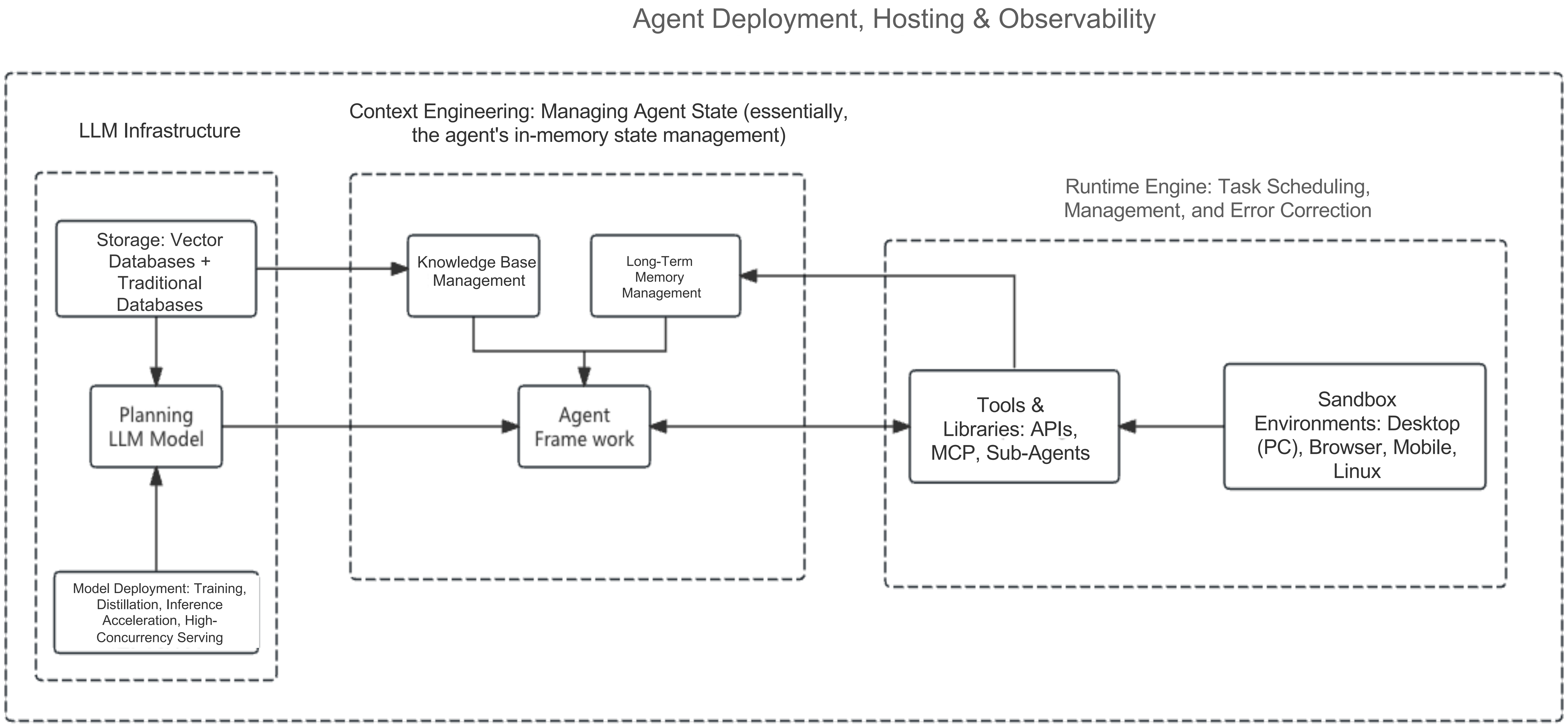
Daytona: An open-source product from Ockam that provides a command-line interface compatible with major IDEs. It allows for one-click startup of cloud sandboxes for code execution, has over 13,000 stars on GitHub, and numerous contributors.

Mobile Use

Droidrun: An open-source Android automation project that gained traction early in the year. It can operate both local and virtual phones and provides a closed-source, end-to-end Mobile Agent service. It raised €2 million in July.

BrowserStack App Automation: This product focuses on multi-scenario, multi-device mobile app testing automation. Its key selling point is a globally deployed resource pool of real devices, compatible with major automation frameworks.

AgentBay's Role in the Agent Application Ecosystem


Agent Deployment, Hosting & Observability
LLM Infrastructure
Storage: Vector Databases + Traditional Databases
Planning LLM Model
Model Deployment: Training, Distillation, Inference Acceleration, High-Concurrency Serving
Context Engineering: Managing Agent State (essentially, the agent's in-memory state management)
Knowledge Base Management
Long-Term Memory Management
Agent Frame work
Runtime Engine: Task Scheduling, Management, and Error Correction
Tools & Libraries: APIs, MCP, Sub-Agents
Sandbox Environments: Desktop (PC), Browser, Mobile, Linux

Modern Agent applications are converging around three core modules:

- Large Language Models (LLMs): Responsible for intent recognition, task decomposition, and multimodal generation. Analogy: the brain 🧠
- Agent frameworks: Responsible for context management, memory, knowledge bases, and agent task orchestration. Analogy: the body's core systems 🫀
- Toolset & execution environment: The service layer that enables agents to perform actions and deliver business value. Analogy: the hands and feet 👋🦶

AgentBay operates at the **toolset & execution environment** layer. This is often the final component customers integrate into their stack. The typical customer journey begins with selecting an LLM and a suitable Agent Framework (such as LangChain or Dify). Only after making these foundational choices do they consider the execution environment and tools.

Common pain points in the agent development

Task Planning

- Logical decomposition
- Step-by-step completion
- Multi-agent collaboration

Long-Term Memory

- Cross-session storage
- Multi-tool interaction
- Personalization

Tool Coordination & Integration

- Richness of toolset
- Seamless switching between tasks
- Data consistency across multiple tools

Context Persistence

- Task progress restorability
- Coherent multi-turn conversations
- Real-time environment awareness

Task Sandboxing

- Isolation and security
- Resource management and performance overhead
- Environment consistency

Key Concepts

Agent

An AI entity capable of perceiving its environment, understanding tasks, making autonomous decisions, and executing actions. Key capabilities include: natural language understanding, task planning, tool use, and multimodal interaction.

Agent Framework

A software framework and development platform for building, deploying, and managing Agents. Core features include multi-agent collaboration, task dispatching, state management, and unified APIs.

Memory (Persistence)

An Agent‘s ability to store and retrieve information, allowing it to maintain state across multiple sessions. This includes user state persistence (config files, cookies, and session data), cross-device data synchronization, and integration with knowledge graphs and vector databases.

Context

The collection of information an Agent requires to execute a task, including session state, environment variables, and historical data. It is essential for maintaining task continuity, enabling state recovery, and optimizing decision-making.

Runtime

The computational environment where an Agent executes its tasks, providing required resources and services. Core components typically include a standardized execution environment, pre-integrated tools, and resource management capabilities.

Tools

External functions or service APIs that an Agent can call to perform specific tasks. Examples include web search, browser control, file I/O, terminal commands, and Python code execution.

Sandbox

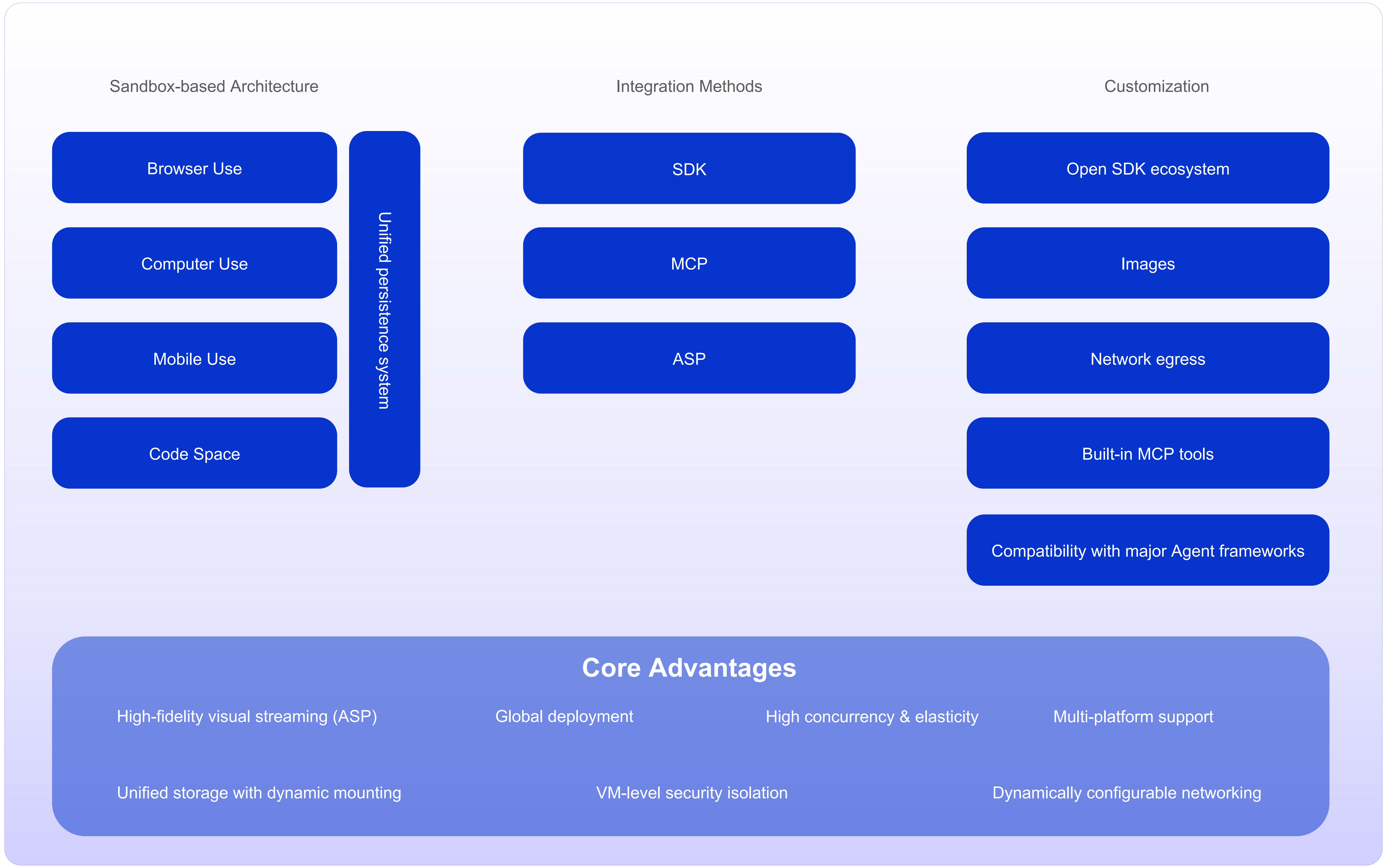
A secure, isolated environment with dedicated compute resources that ensures Agent tasks cannot impact the host system. Key characteristics include: elastic scheduling, complete isolation, automatic reset after each session, zero data retention (non-persistent), and detailed audit trails.

02 / WHAT IS AGENTBAY?

Alibaba Cloud AgentBay: Core Infrastructure for AI Agents

AgentBay is an AI-native infrastructure service (Agent Infra) designed to provide agents with a comprehensive range of execution environments and a suite of integrated tools, along with intelligent agent creation capabilities.

Core Capabilities



Sandbox-based Architecture

- Browser Use
- Computer Use
- Mobile Use
- Code Space
- Unified persistence system

Integration Methods

- SDK
- MCP
- ASP

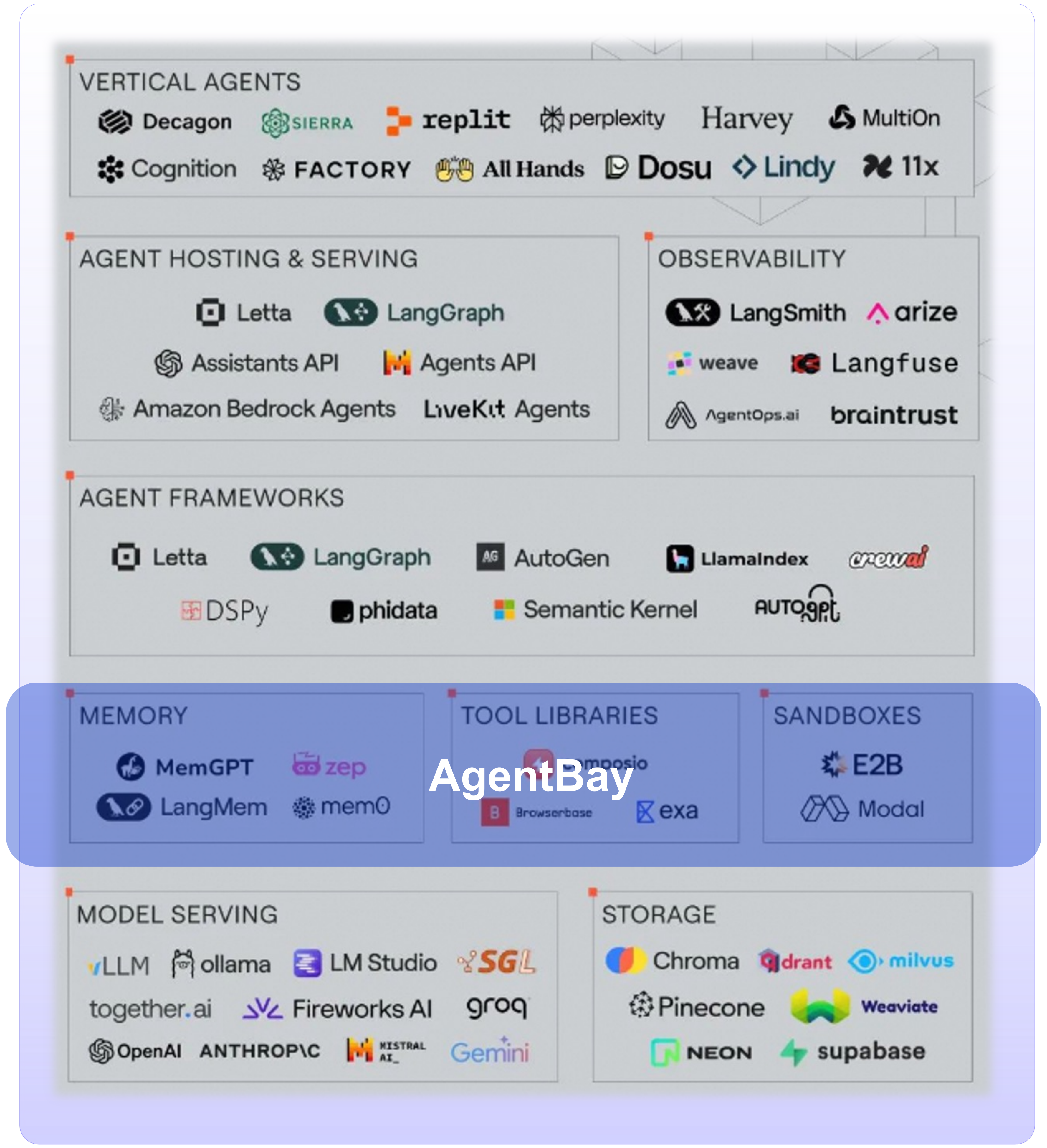
Customization

- Open SDK ecosystem
- Images
- Network egress
- Built-in MCP tools
- Compatibility with major Agent frameworks

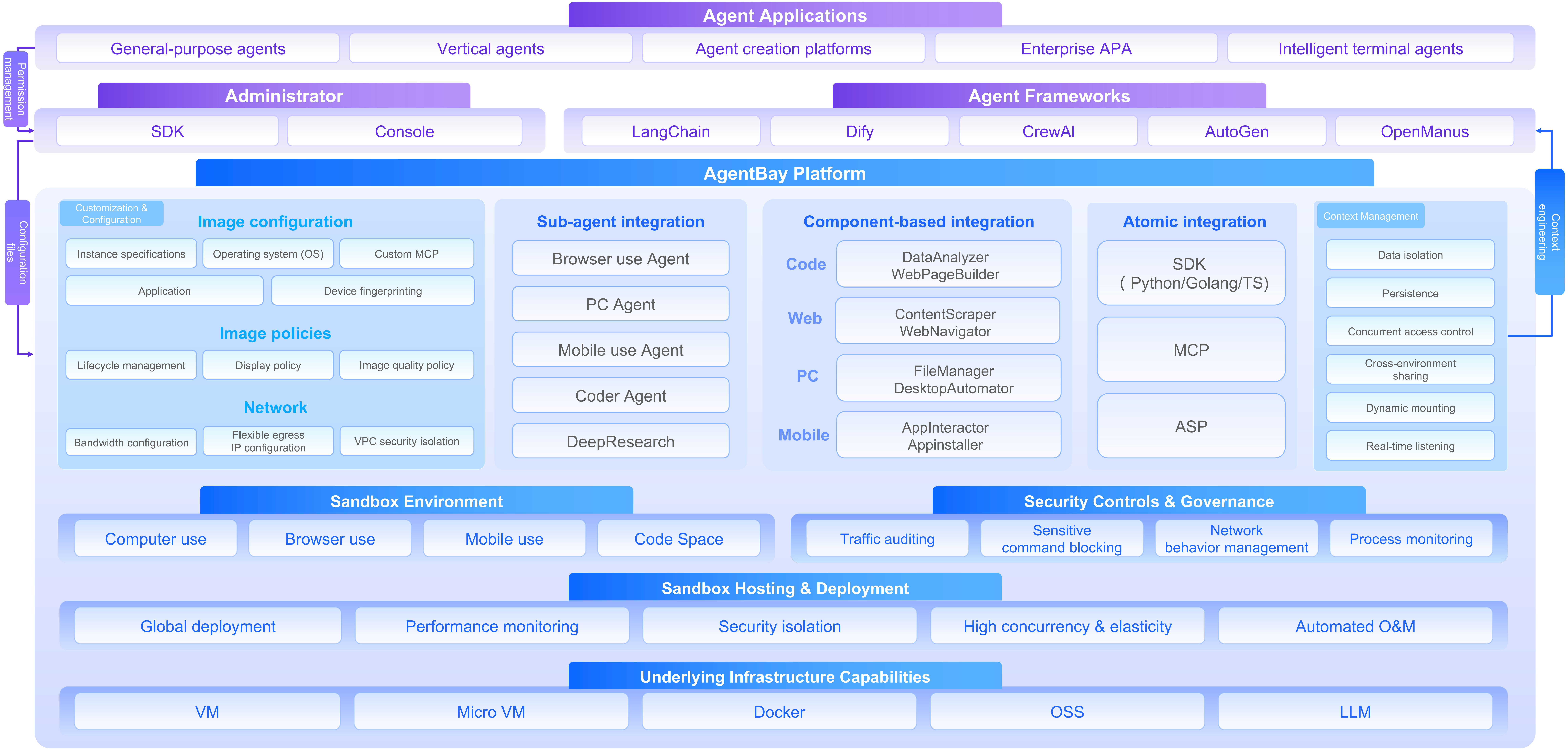
Core Advantages

- High-fidelity visual streaming (ASP)
- Global deployment
- High concurrency & elasticity
- Multi-platform support
- Unified storage with dynamic mounting
- VM-level security isolation
- Dynamically configurable networking

The Agent Ecosystem



AgentBay Architecture Overview



Agent Applications

General-purpose agents | Vertical agents | Agent creation platforms | Enterprise APA | Intelligent terminal agents

Permission management

Administrator

SDK | Console

Agent Frameworks

LangChain | Dify | CrewAI | AutoGen | OpenManus

Context engineering

AgentBay Platform

Configuration files

Customization & Configuration

Image configuration

Instance specifications | Operating system (OS) | Custom MCP

Application | Device fingerprinting

Image policies

Lifecycle management | Display policy | Image quality policy

Network

Bandwidth configuration | Flexible egress IP configuration | VPC security isolation

Sub-agent integration

Browser use Agent

PC Agent

Mobile use Agent

Coder Agent

DeepResearch

Component-based integration

Code: DataAnalyzer, WebPageBuilder

Web: ContentScraper, WebNavigator

PC: FileManager, DesktopAutomator

Mobile: AppInteractor, Appinstaller

Atomic integration

SDK (Python/Golang/TS)

MCP

ASP

Context Management

Data isolation

Persistence

Concurrent access control

Cross-environment sharing

Dynamic mounting

Real-time listening

Sandbox Environment

Computer use | Browser use | Mobile use | Code Space

Security Controls & Governance

Traffic auditing | Sensitive command blocking | Network behavior management | Process monitoring

Sandbox Hosting & Deployment

Global deployment | Performance monitoring | Security isolation | High concurrency & elasticity | Automated O&M

Underlying Infrastructure Capabilities

VM | Micro VM | Docker | OSS | LLM

Core Functions

Browser Use

Drive web automation with multimodal models

Computer Use

Application awareness and control for Linux and Windows

Mobile Use

Intelligent, large-scale application execution

Code Space

Efficient and secure isolated code execution

Unified Persistence System
AgentBay Context

Global Deployment

Built on Alibaba Cloud's infrastructure for a globally distributed network, ensuring low latency, high stability, and a consistent service experience.

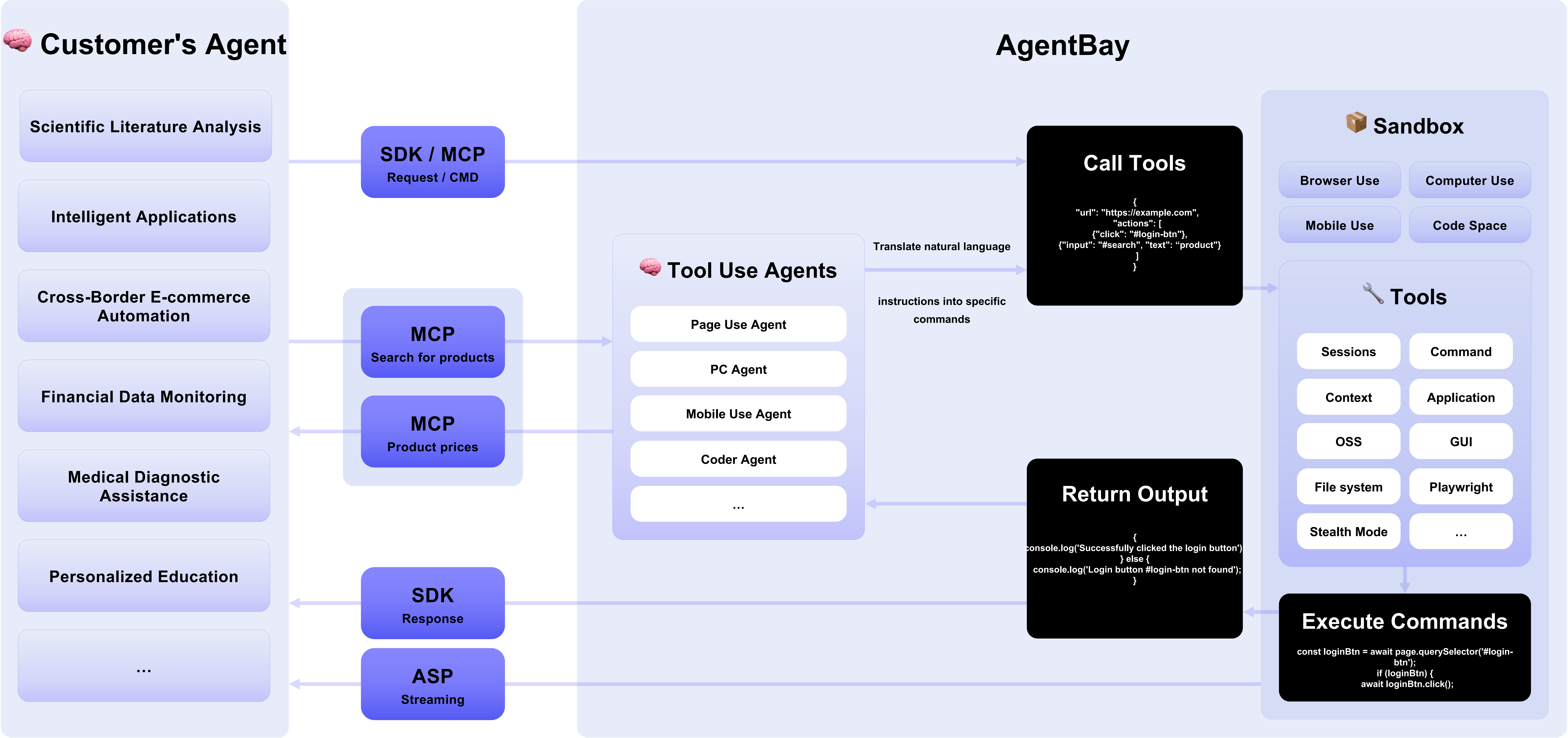
Secure Isolation

Features multi-layered security protection, strict permission isolation, and encrypted data transmission for enterprise-grade security.

Elastic High Concurrency

Supports elastic scaling, dynamic resource allocation, and massive concurrency to handle peak traffic demands

How Customer Agents Use AgentBay to Complete Tasks



Customer's Agent

Scientific Literature Analysis

Intelligent Applications

Cross-Border E-commerce Automation

Financial Data Monitoring

Medical Diagnostic Assistance

Personalized Education

...

SDK / MCP
Request / CMD

MCP
Search for products

MCP
Product prices

SDK
Response

ASP
Streaming

AgentBay

Tool Use Agents

Page Use Agent

PC Agent

Mobile Use Agent

Coder Agent

...

Translate natural language

instructions into specific commands

Call Tools

```
{
"url": "https://example.com",
"actions": [
{"click": "#login-btn"},
{"input": "#search", "text": “product"}
]
}
```

Return Output

```
{
console.log('Successfully clicked the login button');
} else {
console.log('Login button #login-btn not found');
}
```

Sandbox

Browser Use

Computer Use

Mobile Use

Code Space

Tools

Sessions

Command

Context

Application

OSS

GUI

File system

Playwright

Stealth Mode

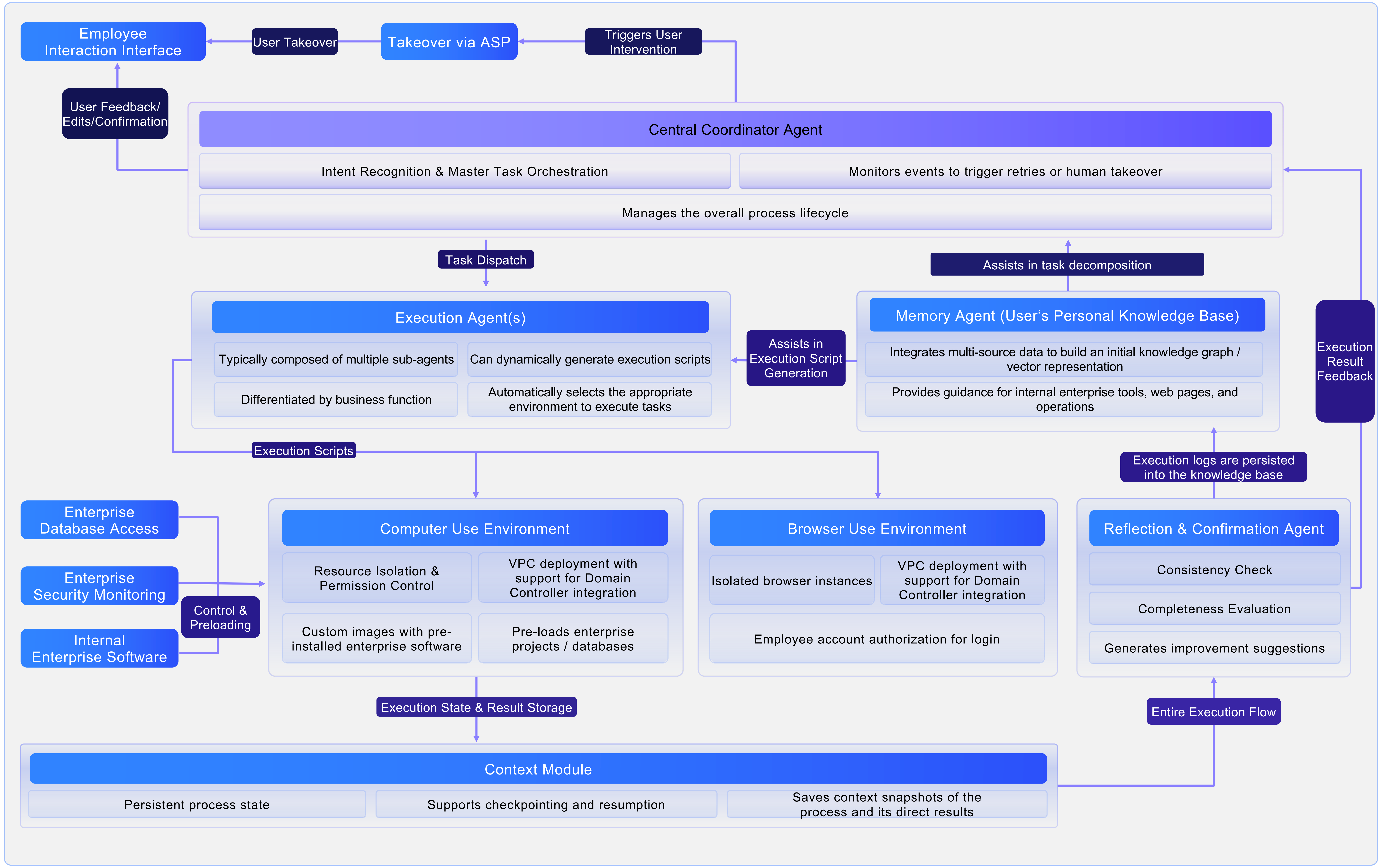
...

Execute Commands

```
const loginBtn = await page.querySelector('#login-
btn');
if (loginBtn) {
await loginBtn.click();
```

Typical Architecture for an Enterprise Process Automation Agent Built on AgentBay

阿里云智能集团
ALIBABA CLOUD INTELLIGENCE GROUP



Employee Interaction Interface

User Takeover

Takeover via ASP

Triggers User Intervention

User Feedback/ Edits/Confirmation

Central Coordinator Agent

Intent Recognition & Master Task Orchestration

Monitors events to trigger retries or human takeover

Manages the overall process lifecycle

Task Dispatch

Assists in task decomposition

Execution Agent(s)

Typically composed of multiple sub-agents

Can dynamically generate execution scripts

Differentiated by business function

Automatically selects the appropriate environment to execute tasks

Assists in Execution Script Generation

Memory Agent (User's Personal Knowledge Base)

Integrates multi-source data to build an initial knowledge graph / vector representation

Provides guidance for internal enterprise tools, web pages, and operations

Execution Result Feedback

Execution Scripts

Execution logs are persisted into the knowledge base

Enterprise Database Access

Enterprise Security Monitoring

Internal Enterprise Software

Control & Preloading

Computer Use Environment

Resource Isolation & Permission Control

VPC deployment with support for Domain Controller integration

Custom images with pre-installed enterprise software

Pre-loads enterprise projects / databases

Browser Use Environment

Isolated browser instances

VPC deployment with support for Domain Controller integration

Employee account authorization for login

Reflection & Confirmation Agent

Consistency Check

Completeness Evaluation

Generates improvement suggestions

Execution State & Result Storage

Entire Execution Flow

Context Module

Persistent process state

Supports checkpointing and resumption

Saves context snapshots of the process and its direct results

1. Central Coordination & Orchestration

The Central Coordinator Agent acts as the "brain," understanding user intent and decomposing complex tasks into executable sub-tasks.

2. Multi-Agent Collaboration

Specialized agents work together, each with a distinct role. Execution Agents perform specific operations, while Memory Agents provide knowledge support.

3. Dynamic Environment Adaptation

The system intelligently selects the appropriate execution environment (e.g., Computer Use, Browser Use) based on the task type and dynamically generates the necessary execution scripts.

AgentBay's value: Provides four core sandbox environments (Browser, Code, Computer, Mobile) with intelligent scheduling.

4. Continuous Knowledge Evolution

The Memory Agent builds a knowledge graph, transforming execution experience into a reusable knowledge base.

5. Self-Reflection and Optimization

The Reflection Agent evaluates execution outcomes, checking for consistency and completeness to generate suggestions for improvement.

6. Persistent Process State

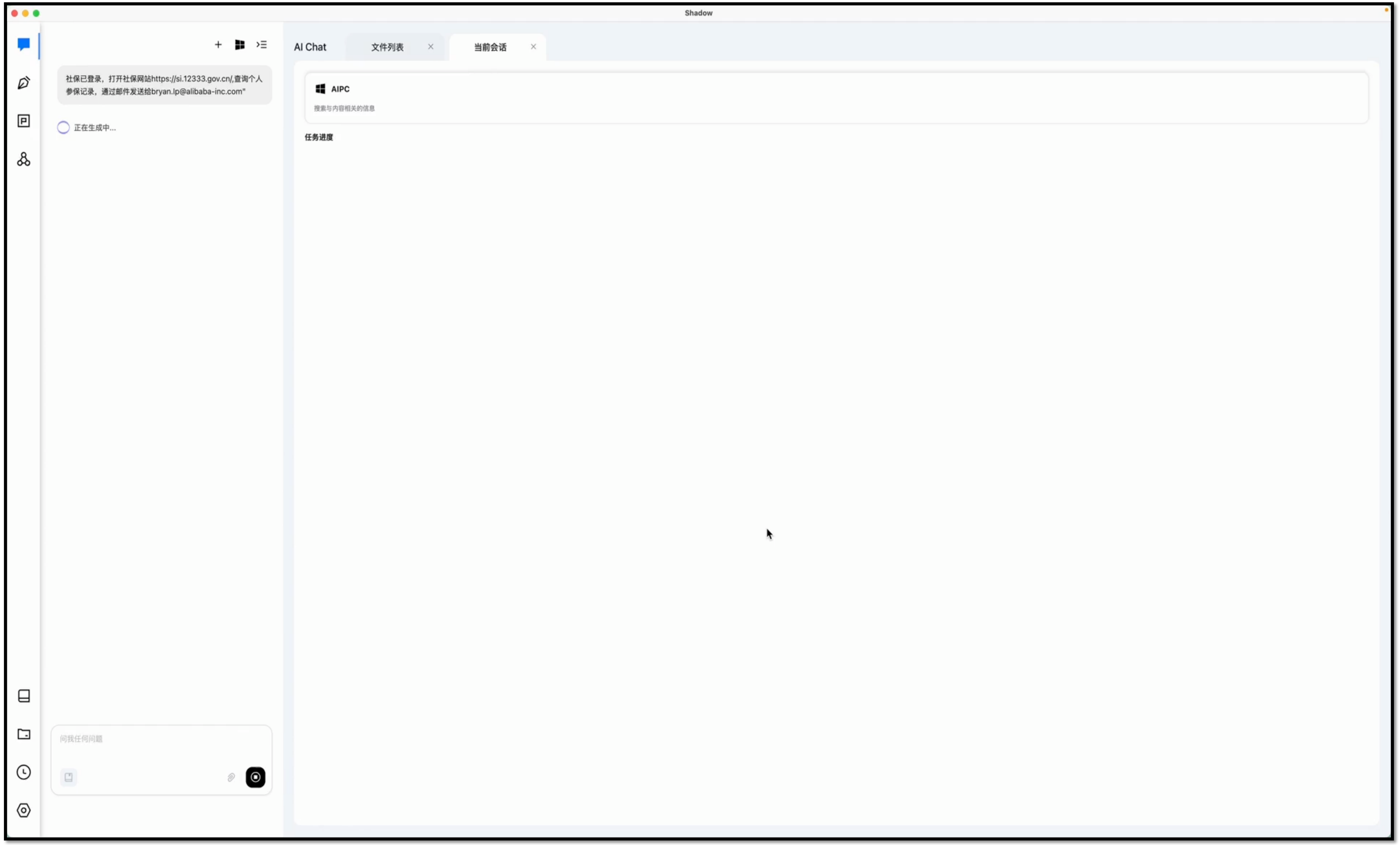
The Context module saves task states, supporting checkpointing and resumption to ensure task reliability.

AgentBay's value: Enables task-related files to persist and be accessible across different environments.

03 / PRODUCT DEMO

Browser Use Demo

When an Agent's task is entirely performed within the browser and does not involve the local system or mobile apps, Browser Use should be selected. Applicable scenarios: web automation operations, automated testing of web applications, etc.

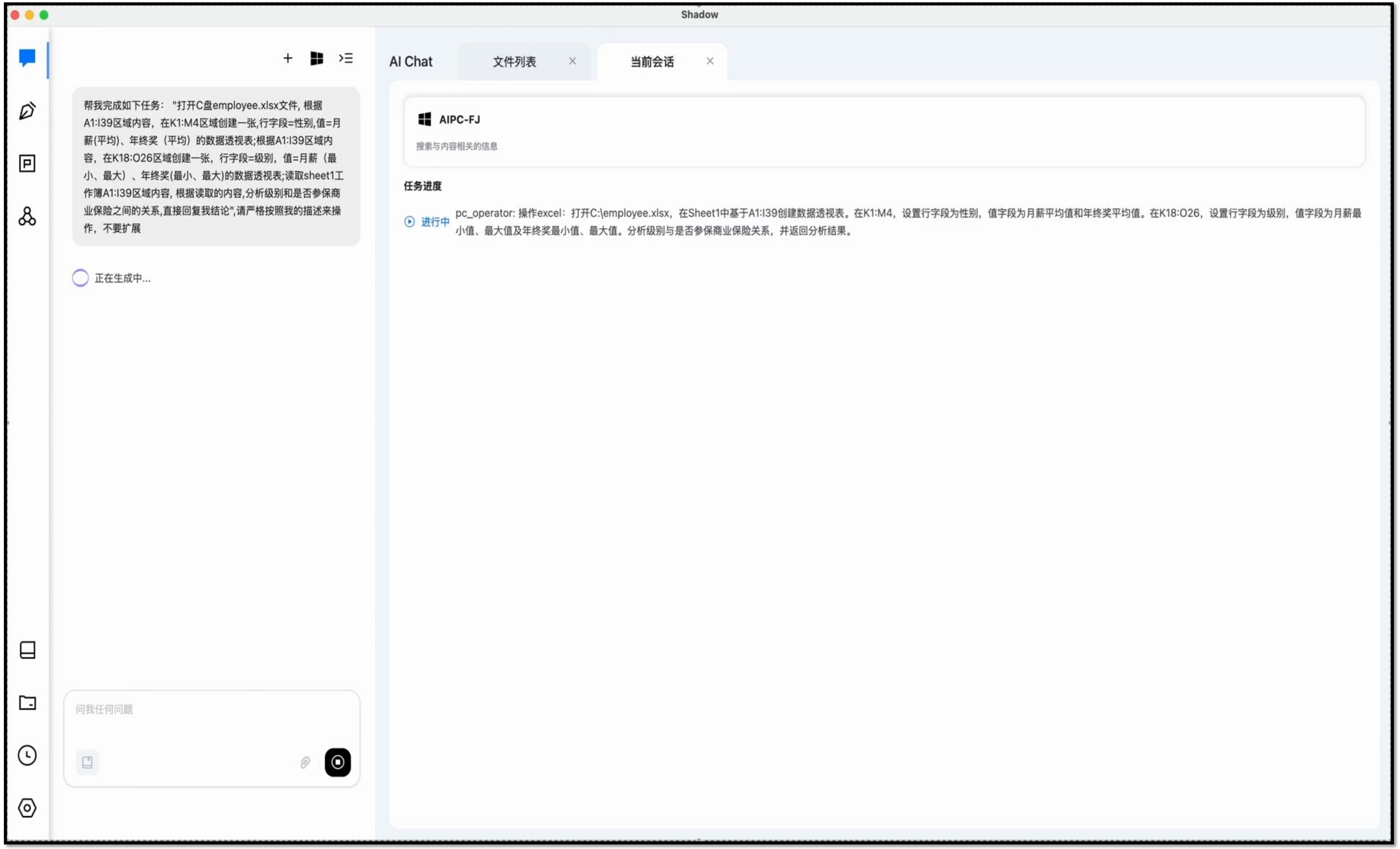


社保已登录，打开社保网站https://si.12333.gov.cn/,查询个人参保记录，通过邮件发送给bryan.lp@alibaba-inc.com"

正在生成中...

Computer Use Demo

When an Agent's task relies on a full operating system and requires installing custom software applications on Windows/Linux systems, Computer Use should be selected. Applicable scenarios: automation of specialized system operations, cross-application data transfer, etc.



帮我完成如下任务："打开C盘employee.xlsx文件, 根据A1:I39区域内容，在K1:M4区域创建一张,行字段=性别,值=月薪(平均)、年终奖（平均）的数据透视表;根据A1:I39区域内容，在K18:O26区域创建一张，行字段=级别，值=月薪（最小、最大）、年终奖(最小、最大)的数据透视表;读取sheet1工作簿A1:I39区域内容, 根据读取的内容,分析级别和是否参保商业保险之间的关系,直接回复我结论",请严格按照我的描述来操作，不要扩展

正在生成中...

任务进度

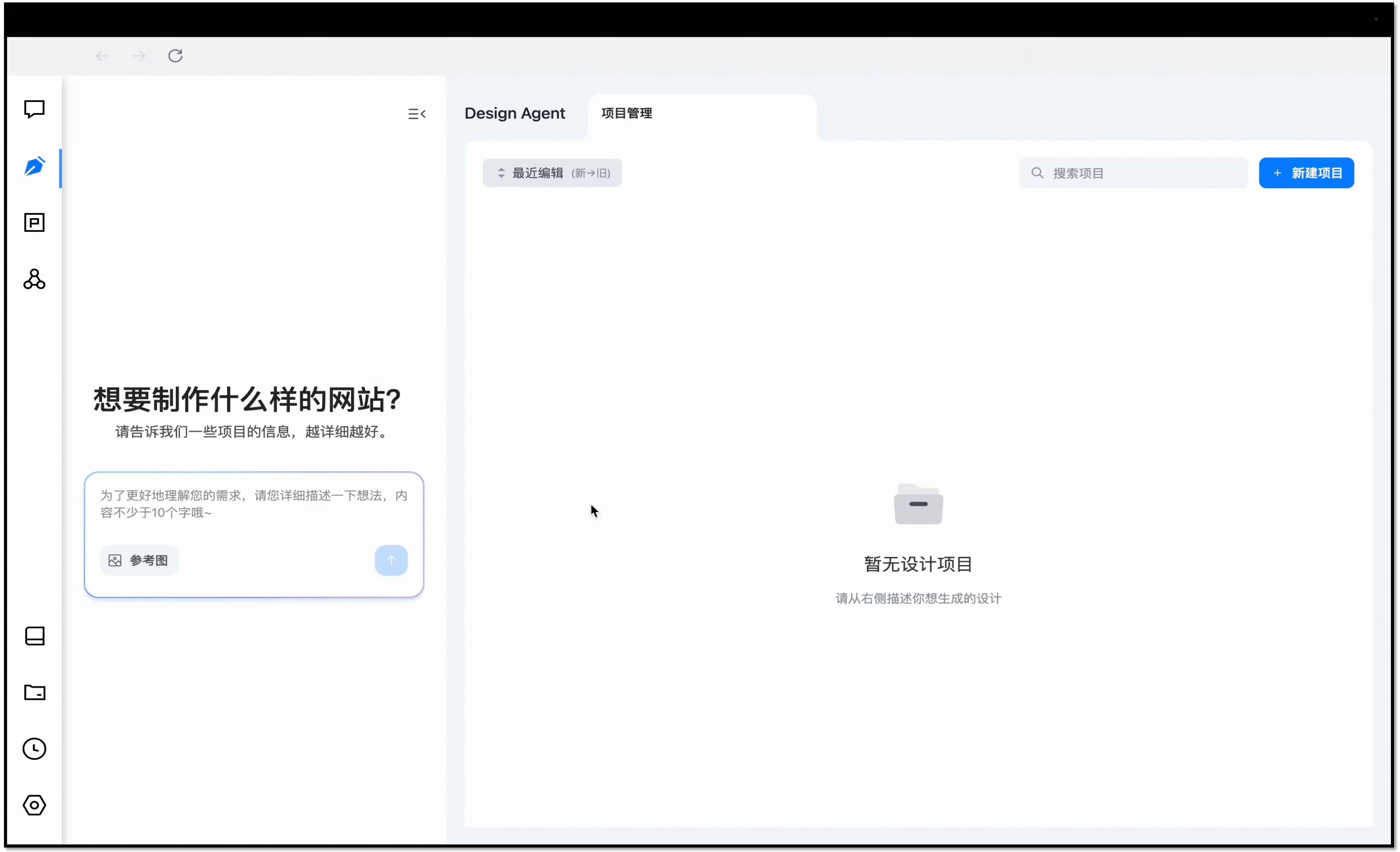
进行中 pc_operator: 操作excel：打开C:\employee.xlsx，在Sheet1中基于A1:I39创建数据透视表。在K1:M4，设置行字段为性别，值字段为月薪平均值和年终奖平均值。在K18:O26，设置行字段为级别，值字段为月薪最小值、最大值及年终奖最小值、最大值。分析级别与是否参保商业保险关系，并返回分析结果。

Mobile Use Demo

When an Agent's task must be executed in an Android phone environment or within a native app, especially for social media operations or mobile testing, Mobile Use should be selected. Applicable scenarios: Android app automated testing, social media matrix management, mobile behavior simulation, etc.

CodeSpace Demo

When an Agent's task is centered around code execution, does not require a graphical interface, and emphasizes scalability and isolation, CodeSpace should be selected. Applicable scenarios: large-scale execution of AI-generated code, code quality inspection, etc.



想要制作什么样的网站?

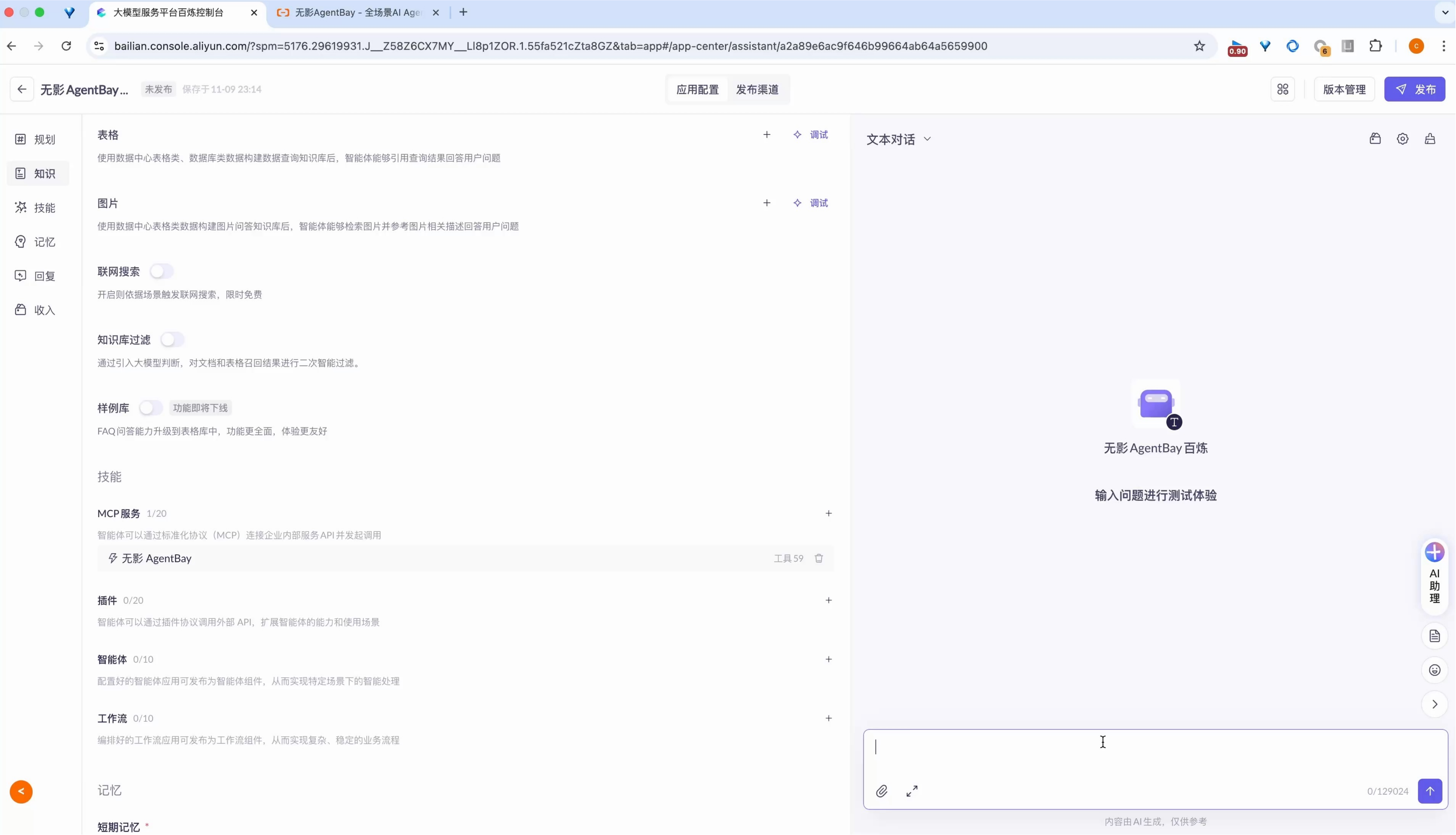
请告诉我们一些项目的信息，越详细越好。

为了更好地理解您的需求，请您详细描述一下想法，内容不少于10个字哦~

暂无设计项目

请从右侧描述你想生成的设计

MCP Server Demo



bailian.console.aliyun.com/?spm=5176.29619931.J__Z58Z6CX7MY__Ll8p1ZOR.1.55fa521cZta8GZ&tab=app#/app-center/assistant/a2a89e6ac9f646b99664ab64a5659900

表格 + 调试

使用数据中心表格类、数据库类数据构建数据查询知识库后，智能体能够引用查询结果回答用户问题

图片 + 调试

使用数据中心表格类数据构建图片问答知识库后，智能体能够检索图片并参考图片相关描述回答用户问题

联网搜索

开启则依据场景触发联网搜索，限时免费

知识库过滤

通过引入大模型判断，对文档和表格召回结果进行二次智能过滤。

样例库 功能即将下线

FAQ 问答能力升级到表格库中，功能更全面，体验更友好

技能

MCP 服务 1/20 +

智能体可以通过标准化协议（MCP）连接企业内部服务 API 并发起调用

无影 AgentBay 工具 59

插件 0/20 +

智能体可以通过插件协议调用外部 API，扩展智能体的能力和使用场景

智能体 0/10 +

配置好的智能体应用可发布为智能体组件，从而实现特定场景下的智能处理

工作流 0/10 +

编排好的工作流应用可发布为工作流组件，从而实现复杂、稳定的业务流程

记忆

短期记忆 *

无影 AgentBay 百炼

输入问题进行测试体验

内容由AI生成，仅供参考

阿里云智能集团
ALIBABA CLOUD INTELLIGENCE GROUP

04 / AGENTBAY SALES GUIDE

Target Customer Profiles

Financial services

Typical use cases:

- Secure and compliant environments for Agent development, testing, and training, with comprehensive audit logs for event tracing.
- Executing Agent tasks in production, such as parsing PDF documents, querying authorized databases, interacting with CRM systems, and sending compliant emails.
- Implementing Human-in-the-Loop (HITL) workflows for exception handling and manual authorization of high-risk operations.

Key requirements:

- Strict adherence to the financial industry‘s legal, regulatory, and compliance standards.
- High-level security for the Agent execution environment to prevent unauthorized access, data leakage, or malicious code execution.
- Achieving quantifiable efficiency gains or operational cost reductions while maintaining security and compliance.

E-commerce

Typical use cases:

- Compliant product information synchronization and price comparison, based on open APIs or authorized data sources.
- Monitoring public opinion on social media and analyzing user feedback.
- Automated data processing for merchants, including generating business reports, querying CRM data, and synchronizing inventory status.

Key requirements:

- Rapidly acquire and process multi-channel product and user data to support decisions on product selection, pricing, and marketing in a highly competitive market.
- Improve the response speed and accuracy of AI applications, such as personalized recommendations and intelligent customer service, through real-time data analysis.

AI applications

Typical use cases:

- General-purpose agent operations within a browser or desktop environment (e.g., automated form filling, cross-application coordination).
- Automating interactions with proprietary GUI applications or game clients.
- A secure, traceable, and high-concurrency code execution environment that supports low-latency, interactive, and risk-controlled task execution.
- Sandbox environment management for Reinforcement Learning (RL) training, supporting state persistence and rapid cloning.

Key requirements:

- Support for large-scale, parallel sandbox environments to meet the compute and isolation requirements of multi-agent reinforcement learning.
- A highly stable, low-latency execution infrastructure to accelerate the productization and technical validation of AI agents.

AgentBay Billing

AgentBay uses a hybrid billing model combining **benefit packages** and **pay-as-you-go** pricing to offer flexible and comprehensive resource options.

Benefit packages

Include concurrent session licenses, management functions, customizations, advanced features, and an included resource allowance.

Tier		Basic	Pro	Ultra
Price		Free + pay-as-you-go for resources	CNY 999/month + pay-as-you-go for resources	CNY 1,499/month + pay-as-you-go for resources
Concurrent sessions	Limit	10	200	200
	Scale-out	—	—	Supported, CNY 10/session/month
Network	Base bandwidth	10 Mbps	10 Mbps	10 Mbps
	Premium bandwidth	—	Supported	Supported
Storage	Included allowance	100 GiB	1 TiB	1 TiB
Images	System images	Supported	Supported	Supported
	Custom images (count)	—	50	50
	Active custom images	—	5	5
Advanced Features	GetLink (Enterprise identity verification)	—	Supported	Supported
	Device fingerprinting	—	—	Supported
	Other	—	—	Early access to new features

Pay-as-you-go

Billed for actual usage of compute, storage, network, and token resources that exceeds the benefit package's included allowance.

Pay-as-you-go		Unit Price	Description
Compute	vCPU	CNY 0.2 /vCPU/hour	Billed per second for actual usage.
	Memory	CNY 0.05 /GiB/hour	Configured at the MiB level with precise measurement. Billed per second.
Tokens	Input tokens	CNY 0.01 / 1,000 tokens	Token consumption from user input when using Alibaba Cloud Workspace's AI services.
	Output tokens	CNY 0.04 / 1,000 tokens	Token consumption from AgentBay's output when using Alibaba Cloud Workspace's AI services.
Network	Premium bandwidth	CNY 0.8 /GiB	Provides dedicated bandwidth with a customizable maximum bandwidth value and a dedicated IP per UID. Billed based on inbound traffic (ingress). SLA is provided.
Storage	Storage	CNY 0.0002 /GiB/hour	For persisting context data and session replay records. Usage beyond the included allowance is billed on a pay-as-you-go basis. Minimum billing unit is MiB. Does not include system/data disks of compute resources.

Package upgrades

- You can upgrade from the Pro tier to the Ultra tier.
- Downgrading from the Ultra tier to the Pro tier is not permitted.
- Upon upgrading, the old benefit package will be immediately billed on a prorated basis for the time used.
- The new package will adopt the expiration date of the previous package, and you will be charged the prorated price difference for the remaining time.

Automatic package downgrades

If a Pro or Ultra package becomes inactive (e.g., due to overdue payments), it will be automatically downgraded to the Basic package. This will result in the following:

- Active sessions using custom images will be terminated immediately.
- Custom image files will be retained for 3 months. If the package is not renewed within this period, the image files will be deleted.
- The included storage allowance will be reduced from 1 TiB to 100 GiB. Any storage usage exceeding 100 GiB will be billed on a pay-as-you-go basis.

Overdue payment rules

If your Alibaba Cloud account has an overdue balance, pay-as-you-go services will be suspended. The following AgentBay features will be impacted:

- You will not be able to create new API Keys. All existing API Keys will be set to an "Overdue" status and become unusable.
- You will not be able to create new custom images. Existing custom images can still be deleted or disabled.
- All custom images will be set to an "Overdue" status and cannot be enabled.
- You will not be able to create new sessions or configure concurrency. Active delivery groups will be released after 24 hours.
- You will not be able to use storage beyond the package's included allowance.

For detailed rules, see the AgentBay documentation at https://www.alibabacloud.com/help/en/agentbay/product-overview/agentbay-billing-instructions

Resources & Links

- SDK (Includes Cookbook, Integration Guides, and Code Samples): https://github.com/aliyun/wuying-agentbay-sdk
- Product Page: https://www.alibabacloud.com/product/agentbay
- Product Documentation: https://www.alibabacloud.com/help/agentbay
- Management Console: https://agentbay.console.aliyun.com/

ALIBABA CLOUD
INTELLIGENCE GROUP

THANKS