

Test 1: Scraping

```
Query: How much does cloudrift ai (https://www.cloudrift.ai/inference) charge for deepseek v3?
2026-01-06 07:57:44,987 - mcp.server.lowlevel.server - INFO - Processing request of type CallToolRequest
2026-01-06 07:57:46,899 - mcp.server.lowLevel.server - INFO - Processing request of type CallToolRequest
2026-01-06 07:57:46,900 - __main__ - INFO - Extracting information for identifier: cloudrift
2026-01-06 07:57:46,901 - __main__ - INFO - Files in scraped_content: ['CloudRift AI.html.txt', 'CloudRift AI_markdown.txt', 'cloudrift_ai_html.txt', 'cloudrift_ai_markdown.txt', 'cloudrift_html.txt', 'cloudrift_markdown.txt', 'deepinfra_html.txt', 'deepinfra_markdown.txt', 'fireworks_html.txt', 'fireworks_markdown.txt', 'groq_html.txt', 'groq_markdown.txt', 'scraped_metadata.json', 'Spring AI_html.txt', 'Spring AI_markdown.txt']
2026-01-06 07:57:46,901 - __main__ - INFO - Checking metadata file: scraped_content\scraped_metadata.json
2026-01-06 07:57:46,902 - __main__ - INFO - Found matching metadata for cloudrift
Based on the information from Cloudrift AI's inference pricing page, **DeepSeek V3 costs**:
```

- **Input**: \$0.15 per 1K tokens
- **Output**: \$0.40 per 1K tokens

The model also has a 163.84K context window available.

```
Query: How much does deepinfra (https://deepinfra.com/pricing) charge for deepseek v3?
2026-01-06 07:58:01,271 - mcp.server.lowlevel.server - INFO - Processing request of type CallToolRequest
2026-01-06 07:58:03,738 - mcp.server.lowLevel.server - INFO - Processing request of type CallToolRequest
2026-01-06 07:58:03,738 - __main__ - INFO - Extracting information for identifier: deepinfra
2026-01-06 07:58:03,739 - __main__ - INFO - Files in scraped_content: ['CloudRift AI.html.txt', 'CloudRift AI_markdown.txt', 'cloudrift_ai_html.txt', 'cloudrift_ai_markdown.txt', 'cloudrift_html.txt', 'cloudrift_markdown.txt', 'deepinfra_html.txt', 'deepinfra_markdown.txt', 'fireworks_html.txt', 'fireworks_markdown.txt', 'groq_html.txt', 'groq_markdown.txt', 'scraped_metadata.json', 'Spring AI_html.txt', 'Spring AI_markdown.txt']
2026-01-06 07:58:03,739 - __main__ - INFO - Checking metadata file: scraped_content\scraped_metadata.json
2026-01-06 07:58:03,739 - __main__ - INFO - Found matching metadata for deepinfra
Based on the Deepinfra pricing page, **DeepSeek V3 costs**:
```

- **Input**: \$0.32 per 1M tokens
- **Output**: \$0.89 per 1M tokens

The model has a 160K context window available.

Test2: Asking a Question

```
Query: Compare cloudrift ai and deepinfra's costs for deepseek v3
Here's a comparison of DeepSeek V3 pricing between Cloudrift AI and Deepinfra:
```

Provider	Input Cost	Output Cost	Context Window
Cloudrift AI*	\$0.15 per 1M tokens	\$0.40 per 1M tokens	163.84K
Deepinfra*	\$0.32 per 1M tokens	\$0.89 per 1M tokens	160K

Summary:

Cloudrift AI is significantly cheaper for DeepSeek V3:
- **Input tokens**: 2.13x cheaper (\$0.15 vs \$0.32)
- **Output tokens**: 2.23x cheaper (\$0.40 vs \$0.89)

Cloudrift AI offers the best value for DeepSeek V3 inference with roughly half the cost of Deepinfra for both input and output tokens.

Test 3: Checking the Database

```
Query: show data

Recently Stored Data:
=====
Pricing Plans:
[{'company_name': 'DeepInfra', 'plan_name': 'DeepSeek-V3', 'input_tokens': 0.32, 'output_tokens': 0.89, 'currency': 'USD'}, {'company_name': 'CloudRift AI', 'plan_name': 'DeepSeek V3', 'input_tokens': 0.15, 'output_tokens': 0.4, 'currency': 'USD'}]
=====
```