Scraper Agent API `0.1.0` `OAS 3.1`

/openapi.json

default ∧

POST `/chat` Chat ∧

Handle a chat request by running the scraping agent, persisting state, and returning a summarized response with structured results.

This endpoint:

- Ensures a session exists and appends the user's prompt to the conversation.
- Loads recent conversation memory (up to 20 messages) and any prior scrape results for the session.
- Invokes the scraping agent with the provided URL, prompt, and limits.
- Saves the agent's outputs (schema, items, visited count) for future reuse.
- Adds a helpful assistant message to the conversation, preferring an LLM-provided summary when the agent reuses previous results.

Args:

```
req (ChatRequest):

    session_id (str): Identifier for the chat session; used to persist memory and results.

    user_prompt (str): The user's instruction or question guiding the agent.

    url (str): The starting URL for the scraper/agent to process.

    max_pages (int | None): Optional page-visit limit for the crawl.

    max_iterations (int | None): Optional iteration cap for the agent.
```

Returns:

```
ChatResponse:

    text (str): A human-readable summary of what was done (reuse summary or scrape summary).

    results (list[dict]): The structured items extracted by the agent.

    session_id (str): Echo of the provided session identifier.

    schema (dict | None): The inferred or provided schema describing item structure.

    visited_count (int | None): Number of pages the agent reports having visited.
```

Side Effects:

- Writes conversation messages and agent outputs to persistent storage for the session.
- Emits log entries describing the run (session, URL, limits).
- May reuse and refine previously saved results depending on the agent's chosen mode.

Raises:

- Propagates I/O, network, validation, or agent execution errors that occur during run or persistence.

Parameters

[Cancel] [Reset]

No parameters

Request body `required`

`application/json` ▾

Edit Value | Schema

```
{
  "url": "https://support.hp.com/us-en/drivers/printers",
  "user_prompt": "Find all printer driver details through the pages.",
  "max_pages": 5,
  "max_iterations": 5,
  "session_id": "1240"
}
```

[Execute] [Clear]

Responses

Curl

```
curl -X 'POST' \
  'http://localhost:8000/chat' \
  -H 'accept: application/json' \
  -H 'Content-Type: application/json' \
  -d '{
  "url": "https://support.hp.com/us-en/drivers/printers",
  "user_prompt": "Find all printer driver details through the pages.",
  "max_pages": 5,
  "max_iterations": 5,
  "session_id": "1240"
}'
```

Request URL

```
http://localhost:8000/chat
```

Server response

Code	Details
200	Response body

```
{
  "text": "Scraped 25 items from https://support.hp.com/us-en/drivers/printers.",
  "results": [
    {
      "id": "HP DeskJet 2700e All-in-One series",
      "name": "HP DeskJet 2700e All-in-One series",
      "version": "Latest",
      "os": "Windows, macOS",
      "url": "/hp-pps-services/hub/navigation?h_client=s-a-r11839-1&h_product=2100187844&h_lang=en&h_cc=us&h_pagetype=s-002&jumpid=in_r11839_us/en/SWDLanding/TopProducts",
      "release_date": "N/A",
      "support_status": "Supported"
    },
    {
      "id": "HP ENVY 5000 All-in-One Printer series",
      "name": "HP ENVY 5000 All-in-One Printer series",
      "version": "Latest",
      "os": "Windows, macOS",
      "url": "/hp-pps-services/hub/navigation?h_client=s-a-r11839-1&h_product=14095644&h_lang=en&h_cc=us&h_pagetype=s-002&jumpid=in_r11839_us/en/SWDLanding/TopProducts",
      "release_date": "N/A",
      "support_status": "Supported"
    },
```

```
        {
            "id": "HP Laser 100 Printer series",
            "name": "HP Laser 100 Printer series",
            "version": "Latest",
            "os": "Windows, macOS",
            "url": "/hp-pps-services/hub/navigation?h_client=s-a-r11839-1&h_product=24494339&h_lang=en&h_cc=us&h_pagetype=s-002&jumpid=in_r11839_us/en/SWDLanding/TopProducts",
            "release_date": "N/A",
```

Download

Response headers

```
content-length: 7646
content-type: application/json
date: Mon,18 Aug 2025 13:29:47 GMT
server: uvicorn
```

Responses

Code	Description	Links
200	Successful Response	*No links*

Media type

application/json

Controls `Accept` header.

Example Value | Schema

```
{
    "text": "string",
    "results": [
        "string"
    ],
    "session_id": "string",
    "schema": {
        "additionalProp1": {}
    },
    "visited_count": 0
}
```

| 422 | Validation Error | *No links* |

Media type

application/json

Example Value | Schema

```
{
    "detail": [
        {
            "loc": [
                "string",
                0
            ],
            "msg": "string",
            "type": "string"
        }
    ]
}
```

Schemas

ChatRequest > Expand all **object**

ChatResponse > Expand all **object**

HTTPValidationError > Expand all **object**

ValidationError > Expand all **object**