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

```json
{
  "url": "https://medrecruit.medworld.com/jobs/list?location=New+South+Wales&page=1",
  "user_prompt": "From the URL, get the role, location, and pay of jobs from page 1 until 3.",
  "max_pages": 3,
  "max_iterations": 3,
  "session_id": "1238"
}
```

[Execute]   [Clear]

Responses

Curl

```
curl -X 'POST' \
  'http://localhost:8000/chat' \
  -H 'accept: application/json' \
  -H 'Content-Type: application/json' \
  -d '{
  "url": "https://medrecruit.medworld.com/jobs/list?location=New+South+Wales&page=1",
  "user_prompt": "From the URL, get the role, location, and pay of jobs from page 1 until 3.",
  "max_pages": 3,
  "max_iterations": 3,
  "session_id": "1238"
}'
```

Request URL

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

Server response

Code	Details
200	**Response body**

```json
{
  "text": "Scraped 114 items from https://medrecruit.medworld.com/jobs/list?location=New+South+Wales&page=1.",
  "results": [
    {
      "role": "Resident Medical Officer",
      "location": "Matiland, New South Wales AU",
      "pay": 140,
      "name": "Emergency Medicine (ED)",
      "error": false
    },
    {
      "role": "Registrar",
      "location": "Bulli, New South Wales AU",
      "pay": 180,
      "name": "Rehabilitation Medicine",
      "error": false
    },
    {
      "role": "Registrar",
      "location": "Coffs Harbour, New South Wales AU",
      "pay": 140,
```

```
        "name": "Geriatric Medicine",
        "error": false
    },
    {
        "role": "Registrar",
        "location": "Ballina, New South Wales AU",
        "pay": 180,
```

Response headers

```
content-length: 14510
content-type: application/json
date: Mon,18 Aug 2025 12:58:13 GMT
server: uvicorn
```

Responses

Code	Description	Links
200	Successful Response	No links

Media type

application/json ▾

Controls Accept header.

Example Value | Schema

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

| 422 | Validation Error | No links |

Media type

application/json ▾

Example Value | Schema

```json
{
    "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**