

# Samahub API Documentation

Samasource provides a REST API as a simple and direct way for clients to integrate with the Samahub. Via the API, clients can programmatically query the status of their projects and jobs, post new jobs to be run, and subscribe to automatically receive job results.

## URLs

The API can be accessed through URLs that begin with:

```
https://hub.samasource.org/api/v1
```

## Authentication and Security

To guarantee data security, clients will only be allowed to connect to the Samahub API using SSL and URLs that begin with “https”. Any request to use the API using an http URL will result in a redirect (status code 302) to a secure URL.

A unique API access key will be generated for every client account. All API requests require that the API access key be appended to the end of the URL, such as:

```
https://hub.samasource.org/api/v1/<...>?access_key=<ACCESS_KEY>
```

In addition to a unique API access key, client access can be further restricted to a single IP address.

## Response formats

Users of the Samahub API can choose to receive responses from the API in one of two formats, JSON or XML, by appending .json or .xml to the end of the URL respectively.

## Request formats

For writing data in CREATE and UPDATE operations, the Samahub API mainly supports 3 formats:

- JSON - requires the Content-Type header to be set to `application/json`
- XML - requires the Content-Type header to be set to `application/xml`
- URL-encoded key-value pairs

Also, direct upload of binary data is supported in a number of formats for certain operations.

## HTTP Response Codes

Depending on the requested operation, the Samahub API will respond with one of the following HTTP status codes:

- 200 - Success. Request was processed successfully.
- 201 - Created. A new resource was created successfully.
- 302 - Redirect. We are redirecting you to a resource. The body of this response often contains a message.
- 400 - Invalid Format. You failed to pass one or more required parameters.
- 401 - Access Denied. Authorization failed, check your key.
- 404 - Resource not found. We could not find the resource (i.e. object) you were looking for, or you do not have access to it.
- 405 - Method Not Allowed. The method you are using (GET, POST, PUT, DELETE) is not available from this URL.
- 406 - Not Acceptable. The format you requested is not available.
- 422 - Unprocessable Entity. There was an error creating a resource.
- 500 - Internal Server Error.

## Errors

In cases where the API responds with an error code, the response body will contain additional information about the error. XML error responses will look like this:

```
<?xml version="1.0" encoding="UTF-8"?>
<errors>
  <error>Access Denied.</error>
</errors>
```

JSON general error responses will look like this:

```
{"errors":["Access Denied."]}
```

Error responses to resource-specific operations will associate error messages to resource attributes, like this:

```
{"errors":{"data":["not provided."]}
```

## Job Data

The Samahub API supports the ability for a client to upload new jobs in a variety of formats. The fundamental component of a Job is task data that is required for work to be done. Within the Samahub, individual tasks are defined as sets of key-value pairs. Because Samahub projects support repeatable work that can happen on an on-going basis, the keys across tasks and even jobs are not expected to change - only the values. During the process of project initiation, clients can work with the Samasource account management team to agree upon a set of task data keys for future incoming jobs. The Samahub API will validate all uploaded jobs so that incoming task data conforms to the schema that was configured for the project. The following examples illustrate how an uploaded job and its task data is expected to be represented in the different formats that the API supports.

## XML:

```
<job>
  <tasks>
    <task>
      <data>
        <name>John Doe</name>
        <age type="integer">20</age>
        <gender>Male</gender>
      </data>
    </task>
    <task>
      <data>
        <name>Jane Doe</name>
        <age type="integer">30</age>
        <gender>Female</gender>
      </data>
    </task>
  </tasks>
</job>
```

## JSON:

```
{
  "job": {
    "tasks": [
      {
        "data": {
          "name": "John Doe",
          "age": 20,
          "gender": "Male"
        }
      },
      {
        "data": {
          "name": "Jane Doe",
          "age": 30,
          "gender": "Female"
        }
      }
    ]
  }
}
```

## Spreadsheets (Excel, OpenOffice):

| name     | age | gender |
|----------|-----|--------|
| John Doe | 20  | Male   |
| Jane Doe | 30  | Female |

## CSV:

```
"name","age","gender"
```

```
"John Doe","20","Male"  
"Jane Doe","30","Female"
```

Note: All spreadsheet and tabular data formats must contain a header row as the first row of the file.

## Receiving Job Results

After a job has been completed by workers on the Samahub and sufficiently reviewed, the API will post the results back to the client via the project's `callback_url` attribute. Clients can supply this URL to Samasource account managers during project initiation, or manage it via the API. When sending results, the Samahub API will initiate an HTTP POST request to the url provided by the client. To ensure successful delivery of job results, clients are advised to make sure that the callback URL they provide can accept incoming requests and, in the case of large result sets, is configured to receive bulk post data.

The Samahub API supports two formats for job results data: XML and JSON. For both formats, the data will be structured identically to the job input formats, with the additional presence of task data key-value pairs to hold workers' results or judgements. Below are examples of job results in XML and JSON, with new worker task data in bold:

XML:

```
<job>  
  <project_id type="integer">16</project_id>  
  <id type="integer">200</id>  
  <state>complete</state>  
  <tasks>  
    <task>  
      <id type="integer">64534</id>  
      <data>  
        <name>John Doe</name>  
        <age type="integer">20</age>  
        <gender>Male</gender>  
      </data>  
      <results>  
        <phone_no>234-555-1234</phone_no>  
        <is_student>YES</is_student>  
      </results>  
    </task>  
    <task>  
      <id type="integer">64535</id>  
      <data>  
        <name>Jane Doe</name>  
        <age type="integer">30</age>  
        <gender>Female</gender>  
      </data>  
      <results>  
        <phone_no>234-555-5678</phone_no>  
        <is_student>NO</is_student>  
      </results>  
    </task>  
  </tasks>
```

```
</job>
```

JSON:

```
{
  "job":
  {
    "project_id":16,
    "id":200,
    "state":"complete",
    "tasks":
    [
      {
        "id": 64534,
        "data":
        {
          "name": "John Doe",
          "age": 20,
          "gender": "Male"
        },
        "results":
        {
          "phone_no": "234-555-1234",
          "is_student": "YES"
        }
      },
      {
        "id": 64535,
        "data":
        {
          "name": "Jane Doe"
          "age": 30,
          "gender": "Female"
        },
        "results":
        {
          "phone_no": "234-555-5678",
          "is_student": "NO"
        }
      }
    ]
  }
}
```

Clients with custom needs for results delivery are encouraged to discuss customized integration solutions with the Samasource account management team.

## API Resources and Methods Reference

### Project

A Project contains the definition for a repeatable specific type of work done by Samasource workers. While Projects will be created and maintained by the Samasource account management team, clients can use the Samahub API to query the projects that are currently set up. Furthermore, clients can use the API to set a project callback URL that the Samahub will use to communicate work results back to the client.

| Attribute    | Type     | Read\Write | Description  |
|--------------|----------|------------|--|
| id           | integer  | R          | Unique identifier for the project.                       |
| title        | string   | R          | A description of the work being done in the project.     |
| callback_url | string   | R\W        | The URL that Samahub will post completed job results to. |
| created_at   | datetime | R          | The date and time at which the project was created.      |

| Method               | Index   |
|----------------------|---|
| Description          | Returns a list of all projects owned by the client.   |
| URL                  | /projects.<FORMAT>?access_key=...   |
| HTTP Method          | GET   |
| Sample XML URL       | https://hub.samasource.org/api/v1/projects.xml?access_key=123   |
| Sample XML response  | <pre>&lt;?xml version="1.0" encoding="UTF-8"?&gt; &lt;projects type="array"&gt;   &lt;project&gt;     &lt;created_at type="datetime"&gt;2011-01-10T15:59:49-08:00&lt;/ created_at&gt;     &lt;title&gt;New Math Challenge&lt;/title&gt;     &lt;callback_url nil="true"&gt;&lt;/callback_url&gt;     &lt;id type="integer"&gt;13&lt;/id&gt;   &lt;/project&gt;   &lt;project&gt;     &lt;created_at type="datetime"&gt;2010-12-16T11:45:52-08:00&lt;/ created_at&gt;     &lt;title&gt;my test project&lt;/title&gt;     &lt;callback_url nil="true"&gt;&lt;/callback_url&gt;     &lt;id type="integer"&gt;12&lt;/id&gt;   &lt;/project&gt;   &lt;project&gt;     &lt;created_at type="datetime"&gt;2010-12-13T12:31:36-08:00&lt;/ created_at&gt;     &lt;title&gt;Vocab reconstruction&lt;/title&gt;     &lt;callback_url nil="true"&gt;&lt;/callback_url&gt;     &lt;id type="integer"&gt;10&lt;/id&gt;   &lt;/project&gt; &lt;/projects&gt;</pre> |
| Sample JSON URL      | https://hub.samasource.org/api/v1/projects.json?access_key=123  |
| Sample JSON response | <pre>[{"project":{"created_at":"2011-01-10T15:59:49-08:00","title":"New Math Challenge","callback_url":null,"id":13}},{"project":{"cre</pre>  |

|                 |   |
|-----------------|---|
|                 | ated_at":"2010-12-16T11:45:52-08:00","title":"my test project","callback_url":null,"id":12}}] |
| Possible Errors | <ul style="list-style-type: none"> <li>● 401 - Bad or missing API access key</li> </ul>       |

|                      |  |
|----------------------|--|
| Method               | <b>Show</b>  |
| Description          | Returns a specific project based on the given id   |
| URL                  | /projects/<ID>.<FORMAT>?access_key=...   |
| HTTP Method          | GET  |
| Sample XML URL       | https://hub.samasource.org/api/v1/projects/16.xml?access_key=123   |
| Sample XML response  | <pre>&lt;?xml version="1.0" encoding="UTF-8"?&gt; &lt;project&gt;   &lt;created_at type="datetime"&gt;2011-02-02T11:16:53-08:00&lt;/created_at&gt;   &lt;title&gt;Business Listing Verification&lt;/title&gt;   &lt;callback_url&gt;http://www.test.com/samahub_results&lt;/callback_url&gt;   &lt;id type="integer"&gt;16&lt;/id&gt; &lt;/project&gt;</pre> |
| Sample JSON URL      | https://hub.samasource.org/api/v1/projects/16.json?access_key=123  |
| Sample JSON response | <pre>{"project":{"created_at":"2011-02-02T11:16:53-08:00","title":"Business Listing Verification","callback_url":"http://www.test.com/samahub_results","id":16}}</pre>   |
| Possible Errors      | <ul style="list-style-type: none"> <li>● 401 - Bad or missing API access key</li> <li>● 404 - requested Project is not found, or not owned by user</li> </ul>  |

|                         |  |
|-------------------------|--|
| Method                  | <b>Update</b>  |
| Description             | Updates writable attributes on a specific project  |
| URL                     | /projects/<ID>.<FORMAT>?access_key=...   |
| HTTP Method             | PUT  |
| Sample XML curl command | curl -X PUT -H 'Content-Type: application/xml' -T <XML_DATA_FILE> https://hub.samasource.org/api/v1/projects/16.xml?access_key=123 |
| Sample XML request data | <pre>&lt;project&gt;   &lt;callback_url&gt;http://www.test.com/samahub_results2&lt;/callback_url&gt; &lt;/project&gt;</pre>        |

|  |   |
|--|---|
| Sample JSON curl command                 | <code>curl -X PUT -H 'Content-Type: application/json -T &lt;JSON_DATA_FILE&gt; https://hub.samasource.org/api/v1/projects/16.json?access_key=123</code>       |
| Sample JSON request data                 | <code>{"project":{"callback_url":"http://www.test.com/samahub_results2"}}</code>  |
| Sample key-value parameters curl command | <code>curl -X PUT -d "project[callback_url]=http://www.test.com/samahub_results2" https://hub.samasource.org/api/v1/projects/16.xml?access_key=123</code>     |
| Possible Errors                          | <ul style="list-style-type: none"> <li>● 401 - Bad or missing API access key</li> <li>● 404 - Requested Project is not found, or not owned by user</li> </ul> |

## Job

A job is a specific batch of work within a project that will be run on the Samahub and delivered back to the client with results. Jobs contain the actual task data required to do work. Clients may use the API to upload new jobs that need to be done, as well as query the status of existing jobs.

| Attribute               | Type     | Read\Write | Description   |
|-------------------------|----------|------------|---|
| id                      | integer  | R          | Unique identifier for the job.  |
| project_id              | integer  | R          | Unique identifier for the project that the job belongs to.  |
| state                   | string   | R          | The current status of the job. Possible values are: "preparation", "in progress", "complete", "canceled", and "error".        |
| processing_error_reason | string   | R          | This attribute will be available for jobs in the 'error' state and will provide a short explanation for the processing error. |
| created_at              | datetime | R          | The date and time at which the job was originally uploaded.   |

| Method      | Index  |
|-------------|--|
| Description | Returns a list of all jobs that belong to the given project_id |

|                      |   |
|----------------------|---|
| URL                  | /projects/<PROJECT_ID>/jobs.<FORMAT>?access_key=...   |
| HTTP Method          | GET   |
| Sample XML URL       | https://hub.samsource.org/api/v1/projects/16/jobs.xml?access_key=123  |
| Sample XML response  | <pre>&lt;?xml version="1.0" encoding="UTF-8"?&gt; &lt;jobs type="array"&gt;   &lt;job&gt;     &lt;project_id type="integer"&gt;16&lt;/project_id&gt;     &lt;id type="integer"&gt;197&lt;/id&gt;     &lt;state&gt;in progress&lt;/state&gt;   &lt;/job&gt;   &lt;job&gt;     &lt;project_id type="integer"&gt;16&lt;/project_id&gt;     &lt;id type="integer"&gt;202&lt;/id&gt;     &lt;state&gt;preparation&lt;/state&gt;   &lt;/job&gt; &lt;/jobs&gt;</pre> |
| Sample JSON URL      | https://hub.samsource.com/api/v1/projects/16/jobs.json?access_key=123   |
| Sample JSON response | [{"job":{"project_id":16,"id":197,"state":"in progress"}}, {"job":{"project_id":16,"id":202,"state":"preparation"}}]  |
| Possible Errors      | <ul style="list-style-type: none"> <li>• 401 - Bad or missing API access key</li> <li>• 404 - Requested Project not found, or not owned by user</li> </ul>  |

|                     |   |
|---------------------|---|
| Method              | <b>Show</b>   |
| Description         | Returns a specific job based on the given id and project_id   |
| URL                 | /projects/<PROJECT_ID>/jobs/<ID>.<FORMAT>?access_key=...  |
| HTTP Method         | GET   |
| Sample XML URL      | https://hub.samsource.com/api/v1/projects/16/jobs/197.xml?access_key=123  |
| Sample XML response | <pre>&lt;?xml version="1.0" encoding="UTF-8"?&gt; &lt;job&gt;   &lt;created_at type="datetime"&gt;2011-02-02T11:18:39-08:00&lt;/created_at&gt;   &lt;project_id type="integer"&gt;16&lt;/project_id&gt;   &lt;id type="integer"&gt;197&lt;/id&gt;   &lt;state&gt;in progress&lt;/state&gt; &lt;/job&gt;</pre> |

|                      |  |
|----------------------|--|
| Sample JSON URL      | <a href="https://hub.samsource.com/api/v1/projects/16/jobs.json?access_key=123">https://hub.samsource.com/api/v1/projects/16/jobs.json?access_key=123</a>  |
| Sample JSON response | <pre>{"job":{"created_at":"2011-02-02T11:18:39-08:00","project_id":16,"id":197,"state":"in progress"}}</pre>   |
| Possible Errors      | <ul style="list-style-type: none"> <li>• 401 - Bad or missing API access key</li> <li>• 404 - Requested Job not found, or not owned by user, or does not belong to the Project that was given</li> </ul> |

|                                |   |
|--------------------------------|---|
| Method                         | <b>Create</b>   |
| Description                    | Uploads a new job to the Samahub  |
| URL                            | /projects/<PROJECT_ID>/jobs.<FORMAT>?access_key=...   |
| HTTP Method                    | POST  |
| Sample XML curl command        | <pre>curl -X POST -H 'Content-Type: application/xml' -T &lt;XML_DATA_FILE&gt; https://hub.samsource.org/api/v1/projects/16/jobs.xml?access_key=123</pre>  |
| Sample XML request data        | <pre>&lt;job&gt;   &lt;tasks&gt;     &lt;task&gt;       &lt;data&gt;         &lt;name&gt;Richards, Denise&lt;/name&gt;       &lt;/data&gt;     &lt;/task&gt;     &lt;task&gt;       &lt;data&gt;         &lt;name&gt;Rodman, Dennis&lt;/name&gt;       &lt;/data&gt;     &lt;/task&gt;     &lt;task&gt;       &lt;data&gt;         &lt;name&gt;Obama, Barack&lt;/name&gt;       &lt;/data&gt;     &lt;/task&gt;     &lt;task&gt;       &lt;data&gt;         &lt;name&gt;Lewis, Jerry&lt;/name&gt;       &lt;/data&gt;     &lt;/task&gt;   &lt;/tasks&gt; &lt;/job&gt;</pre> |
| Sample successful XML response | <pre>&lt;?xml version="1.0" encoding="UTF-8"?&gt; &lt;job&gt;   &lt;project_id type="integer"&gt;16&lt;/project_id&gt;   &lt;id type="integer"&gt;224&lt;/id&gt;   &lt;state&gt;preparation&lt;/state&gt;</pre>   |

|  |   |
|--|---|
|  | </job>  |
| Sample JSON curl command               | curl -X POST -H 'Content-Type: application/json' -T <JSON_DATA_FILE> https://hub.samasource.org/api/v1/projects/16/jobs.json?access_key=123   |
| Sample JSON request data               | <pre>{   "job":   {     "tasks":     [       {         "data":         {           "name": "Albert Einstein"         }       },       {         "data":         {           "name": "Jack White"         }       },       {         "data":         {           "name": "Meg White"         }       }     ]   } }</pre> |
| Sample successful JSON response        | {"job":{"project_id":16,"id":226,"state":"preparation"}}  |
| Sample direct file upload curl command | curl -X POST -H 'Content-Type: <SUPPORTED_CONTENT_TYPE>' -T <DATA_FILE> https://hub.samasource.org/api/v1/projects/16/jobs.json?access_key=123  |
| Supported direct upload content types  | <ul style="list-style-type: none"> <li>● application/vnd.oasis.opendocument.spreadsheet</li> <li>● application/vnd.openxmlformats-officedocument.spreadsheetml.sheet</li> <li>● application/vnd.ms-excel</li> <li>● application/xls</li> <li>● text/csv</li> <li>● text/tab-separated-values</li> </ul>                 |
| Possible Errors                        | <ul style="list-style-type: none"> <li>● 401 - Bad or missing API key</li> <li>● 404 - The Project given by project_id does not exist, or the user does not own it.</li> <li>● 422 - The new Job could not be created because uploaded data is invalid or missing.</li> </ul>   |