

Cosmosid Scraping Data v1

[Base URL: 0.0.0.0:4501 /]
http://0.0.0.0:4501/api/swagger/?format=openapi

API documentation for viewing scraped data

[Contact the developer](#)

Schemes

HTTP

[Django Login](#)[Authorize](#)

Filter by tag

apis

[GET /apis/collected-links/](#)

apis_collected-links_list

[Cancel](#)

Parameters

Name

Description

job_id

string
(query)

job_id

last_updated

last_updated

string
(query)

last_updated

[Execute](#)[Clear](#)

Responses

Response content type

application/json

curl

```
curl -X 'GET' \
  'http://0.0.0.0:4501/apis/collected-links/' \
  -H 'accept: application/json' \
  -H 'X-CSRFToken: jYsBDQs5TevSTfz9gCSt3inPk4nShYAV35jH3Pu37d6GbIJWlw4RxpYLr4W9B6b5'
```



Request URL

[http://0.0.0.0:4501/apis/collected-links/](#)

Server response

Code

Details

200

Response body

```
[{"id": 1, "name": "SRR17798759", "url": "/samples/19a62107-f9fe-4622-92b3-25b97da67c15", "folder": "Urinary Tract Infection Microbiome - Neugent et al 2022 - PRJNA801448 - KEPLER", "last_updated": "2024-10-26T07:02:31.798407Z", "job_id": null}, {"id": 2, "name": "SRR17798707", "url": "/samples/14f34828-aaa1-4609-9fc9-8203b951b26a", "folder": "Urinary Tract Infection Microbiome - Neugent et al 2022 - PRJNA801448 - KEPLER", "last_updated": "2024-10-26T07:02:31.798407Z", "job_id": null}, {"id": 3, "name": "SRR17798719", "url": "/samples/191df5cb-528a-44f5-8d3e-1bc77fb234a3", "folder": "Urinary Tract Infection Microbiome - Neugent et al 2022 - PRJNA801448 - KEPLER", "last_updated": "2024-10-26T07:02:31.798407Z", "job_id": null}, {"id": 4, "name": "SRR17798728",
```

[Download](#)

Response headers

```
allow: GET,POST,HEAD,OPTIONS
content-length: 19766
content-type: application/json
date: Sat, 26 Oct 2024 19:50:19 GMT
referrer-policy: same-origin
server: WSGIServer/0.2 CPython/3.8.10
vary: Accept
x-content-type-options: nosniff
x-frame-options: DENY
```

Request duration

74 ms

Responses

Code

Description

200

Example Value | Model

```
▼ [CollectedLinks ▼
  id          Id ▼ integer
            title: Id
            readOnly: true
  name*      Name ▼ string
            title: Name
            minLength: 1
  url*       Url ▼ string($uri)
            title: Url
            maxLength: 200
            minLength: 1
  folder     Folder ▼ string
            title: Folder
            x-nullable: true
  last_updated LastUpdated ▼ string($date-time)
            title: Last updated
            readOnly: true
  job_id     JobId ▼ string
            title: Job_id
```

x-nullable: true

}]

POST /apis/collected-links/ apis_collected-links_create ↴ 🔒

GET /apis/collected-links/{id}/ apis_collected-links_read ↴ 🔒

Parameters

Name Description

id * required A unique integer value identifying this collected links.
integer (path) 35

Cancel

Execute Clear

Responses Response content type application/json

Curl

```
curl -X 'GET' \
'http://0.0.0.0:4501/apis/collected-links/35/' \
-H 'accept: application/json' \
-H 'X-CSRF_TOKEN: jYSBDQs5TevStfzgCst3inFk4nShYAV35jH3Pu37d6GbIJWLw4RpYLr4W9B6b5'
```

Request URL

```
http://0.0.0.0:4501/apis/collected-links/35/
```

Server response

Code Details

200 Response body

```
{
  "id": 35,
  "name": "SRR2005652_2_0_R-ear_0_20",
  "url": "/samples/e50406f0-636f-4b94-a753-28dce66beba1",
  "folder": "R-ear",
  "last_updated": "2024-10-26T07:02:31.798407Z",
  "job_id": null
}
```

Download

Response headers

```
allow: GET,PUT,PATCH,DELETE,HEAD,OPTIONS
content-length: 174
content-type: application/json
cross-origin-opener-policy: same-origin
date: Sat, 26 Oct 2024 19:21:13 GMT
referrer-policy: same-origin
server: WSGIServer/0.2 CPython/3.8.10
vary: Accept
x-content-type-options: nosniff
x-frame-options: DENY
```

Request duration

108 ms

Responses

Code Description

200 Example Value | Model

```
CollectedLinks ↴ {
  id          integer      title: Id      readOnly: true
  name*       string       title: Name
  url*        string(suri) title: Url      maxLength: 200
  folder       string       title: Folder  x-nullable: true
  last_updated string($date-time) title: Last updated
  job_id      string       title: Job id   x-nullable: true
}
```

PUT /apis/collected-links/{id}/ apis_collected-links_update ↴ 🔒

PATCH /apis/collected-links/{id}/ apis_collected-links_partial_update ↴ 🔒

DELETE /apis/collected-links/{id}/ apis_collected-links_delete ↴ 🔒

GET /apis/exported-results/ apis_exported-results_list ↴ 🔒

Parameters

Name Description

collected_link_job_id collected_link_job_id
string (query) collected_link_job_id

taxonomy_level taxonomy_level
string (query) taxonomy_level

last_updated last_updated
string (query) last_updated

Cancel

Execute Clear

Responses Response content type application/json

Curl

```
curl -X 'GET' \
'http://0.0.0.0:4501/apis/exported-results/' \
-H 'accept: application/json' \
-H 'X-CSRFToken: jYSBDQs5TevStz9gCst3inFk4nShYAV35jH3Pu37d6GbIJLw4RxpYLr4W9B6b5'
```

Request URL

```
http://0.0.0.0:4501/apis/exported-results/
```

Server response

Code	Details
200	<p>Response body</p> <pre>"result": "Bacteria", "taxonomic_level": "phylum", "downloaded_file": "/home/tlabib/Documents/github/cosmosid-scraping-api-zaag/downloads/SRR17798706_kepler-biom_phylum_2024-10-26_18_13 (1).tsv", "last_updated": "2024-10-26T18:13:41.172950Z", "job_id": "272c50d9-fba9-4c3b-a170-a72b38f4290c", }, { "id": 643, "collected_link": { "id": 6, "name": "SRR17798706", "url": "/samples/92c1429a-fdfb-41b0-a7e3-5223b57d8f3d", "folder": "Urinary Tract Infection Microbiome - Neugent et al 2022 - PRJNA801448 - KEPLER", "last_updated": "2024-10-26T07:02:31.798407Z", "job_id": null }, "url": "/samples/92c1429a-fdfb-41b0-a7e3-5223b57d8f3d", "result": "Bacteria", "taxonomic_level": "class", "downloaded_file": "/home/tlabib/Documents/github/cosmosid-scraping-api-zaag/downloads/SRR17798706_kepler-biom_class_2024-10-26_18_13 (1).tsv", "last_updated": "2024-10-26T18:13:46.651910Z", "job_id": "272c50d9-fba9-4c3b-a170-a72b38f4290c", }, { "id": 644, "collected_link": { "id": 6, "name": "SRR17798706", "url": "/samples/92c1429a-fdfb-41b0-a7e3-5223b57d8f3d" }</pre> <p>Download</p>
	Response headers
	<pre>allow: GET,POST,HEAD,OPTIONS content-length: 358412 content-type: application/json cross-origin: true date: Sat, 26 Oct 2024 19:50:26 GMT referrer-policy: same-origin server: WSGIServer/0.2 CPython/3.8.10 vary: Accept x-content-type-options: nosniff x-frame-options: DENY</pre>
	Request duration
	405 ms

Responses

Code	Description
200	<p>Example Value Model</p> <pre>▼ [!ExportedResults ▼ { id ID ▶ integer title ID readOnly true collected_link* CollectedLinks ▶ { id Id > [...] name* Name > [...] url* Url > [...] folder Folder > [...] last_updated Last updated > [...] job_id Job id > [...] } url* Url ▶ string(\$uri) title Url maxLength 200 minLength 1 result* Result ▶ string title Result minLength 1 taxonomy_level* TaxonomyLevel ▶ string title Taxonomy level x-nullable true downloaded_file DownloadedFile ▶ string title Downloaded file x-nullable true last_updated LastUpdated ▶ string(\$date-time) title Last updated readOnly true job_id JobId ▶ string title Job id x-nullable true }]</pre>

POST /apis/exported-results/ apis_exported-results_create

GET /apis/exported-results/{id}/ apis_exported-results_read

Parameters

Name	Description
id * required	A unique integer value identifying this exported results.
integer (path)	5

Execute Clear

Responses Response content type application/json

Curl

```
curl -X 'GET' \
'http://0.0.0.0:4501/apis/exported-results/5/' \
-H 'accept: application/json' \
-H 'X-CSRFToken: jYSBDQs5TevStz9gCst3inFk4nShYAV35jH3Pu37d6GbIJLw4RxpYLr4W9B6b5'
```

Request URL

```
http://0.0.0.0:4501/apis/exported-results/5/
```

Server response

Code	Details
------	---------

200 Response body

```
{
  "id": 5,
  "collected_link": {
    "id": 69,
    "name": "SRR17798746",
    "url": "/samples/bb228d07-8e89-4fab-a742-81a046438f69",
    "folder": "Urinary Tract Infection Microbiome - Neugent et al 2022 - PRJNA801448 - KEPLER",
    "last_updated": "2024-10-26T07:02:31.798407Z",
    "job_id": null
  },
  "url": "/samples/bb228d07-8e89-4fab-a742-81a046438f69",
  "result": "Bacteria",
  "taxonomy_level": "class",
  "downloaded_file": "/home/tlabin/Documents/github/cosmosis-scraping-api-zag/downloads/SRR17798746_kepler-biom_class_2024-10-20_13_52.tsv",
  "last_updated": "2024-10-26T07:02:31.803052Z",
  "job_id": null
}
```

[Download](#)

Response headers

```
allow: GET,PUT,PATCH,DELETE,HEAD,OPTIONS
content-length: 554
content-type: application/json
cross-origin-opener-policy: same-origin
date: Sat, 27 Oct 2024 19:32:28 GMT
referrer-policy: same-origin
server: WSGIServer/0.2 CPython/3.8.10
vary: Accept
x-content-type-options: nosniff
x-frame-options: DENY
```

Request duration

96 ms

Responses

Code	Description
200	Example Value Model

```
ExportedResults ▾ {
  id          integer
              title: ID
              readOnly: true
  collected_link* CollectedLinks ▾ {
    id          integer
              title: Id
              readOnly: true
    name*      Name ▾ string
              title: Name
              minLength: 1
    url*       Url ▾ string($uri)
              title: Url
              maxLength: 200
              minLength: 1
    folder     Folder ▾ string
              title: Folder
              x-nullable: true
    last_updated Last updated ▾ string($date-time)
              title: Last updated
              readOnly: true
    job_id     Job id ▾ string
              title: Job id
              x-nullable: true
  }
  url*       string($uri)
              title: Url
              maxLength: 200
              minLength: 1
  result*    Result ▾ string
              title: Result
              minLength: 1
  taxonomy_level* Taxonomy level ▾ string
              title: Taxonomy level
              x-nullable: true
  downloaded_file string
              title: Downloaded file
              x-nullable: true
  last_updated Last updated ▾ string($date-time)
              title: Last updated
              readOnly: true
  job_id     Job id ▾ string
              title: Job id
              x-nullable: true
}
}
```

PUT /apis/exported-results/{id}/ apis_exported-results_update

PATCH /apis/exported-results/{id}/ apis_exported-results_partial_update

DELETE /apis/exported-results/{id}/ apis_exported-results_delete

GET /apis/exported-results/{id}/tsv_content/ apis_exported-results_tsv_content

Parameters

Name	Description
id <small>required</small>	A unique integer value identifying this exported results.
integer (path)	1

Execute Clear

Responses

Response content type: application/json

Curl

```
curl -X 'GET' \
  http://0.0.0.0:4501/apis/exported-results/1/tsv_content/' \
  -H 'accept: application/json' \
  -H 'X-CSRFToken: jYsBDQs5TeVStfz9qC5t3inFk4nShYAV35jh3Pu37d6Gb1JWlw4RxpYlr4W9B6b5'
```

Request URL

http://0.0.0.0:4501/apis/exported-results/1/tsv_content/

Server response

Code	Details
200	Response body

```
"Tax ID": "2702",
"GTDB ID": "GCA_013315195.1",
"Relative Abundance": "1.41",
"% Total Matches": "0.88757",
"% Unique Matches": "0.138801",
"Reads Frequency": "78977",
"Normalized Reads Frequency": "102789"
},
```

```

    "Name": "S. limosilactobacillus portuensis GCA_902362445.1",
    "Tax ID": "2702",
    "GTDB ID": "GCA_902362445.1",
    "Relative Abundance": "0.79",
    "% Total Matches": "0.563338",
    "% Unique Matches": "0.563338",
    "Reads Frequency": "50613",
    "Normalized Reads Frequency": "57591"
},
{
    "Name": "Bifidobacterium swidsinskii GCA_013315145.1",
    "Tax ID": "2702",
    "GTDB ID": "GCA_013315145.1",
    "Relative Abundance": "0.2299999999999999",
    "% Total Matches": "0.371952",
    "% Unique Matches": "0.010329",
    "Reads Frequency": "12401",
    "Normalized Reads Frequency": "16767"
},

```

Response headers

```

allow: GET, HEAD, OPTIONS
content-length: 3723
content-type: application/json
cross-origin-opener-policy: same-origin
date: Sat, 26 Oct 2024 19:21:45 GMT
referrer-policy: same-origin
server: WSGIServer/0.2 CPython/3.8.10
vary: Accept
x-content-type-options: nosniff
x-frame-options: DENY

```

Request duration

1261 ms

Responses

Code	Description
200	TSV file content
404	File not found
500	Internal server error

GET /apis/scraping-job/job_status/ apis_scraping-job_job_status ^

Parameters

Name	Description
job_id	Job ID to check status
string (query)	4e67d5d2-ca5e-44d6-93cf-60bf3e1815eb

Execute Clear

Responses Response content type application/json

Curl

```
curl -X 'GET' \
  'http://0.0.0.4501/apis/scraping-job/job_status/?job_id=4e67d5d2-ca5e-44d6-93cf-60bf3e1815eb' \
  -H 'accept: application/json' \
  -H 'X-CSRF_TOKEN: jYSBDQs5Tev5Tfz9gCst3inFk4nShYAV35jH3Pu37d6GbIjWlW4RpYlr4W9B6b5'
```

Request URL

http://0.0.0.4501/apis/scraping-job/job_status/?job_id=4e67d5d2-ca5e-44d6-93cf-60bf3e1815eb

Server response

Code Details

Code	Details
200	Response body
	{ "job_id": "4e67d5d2-ca5e-44d6-93cf-60bf3e1815eb", "status": "Completed" }

Response headers

```

allow: GET, HEAD, OPTIONS
content-type: application/json
cross-origin-opener-policy: same-origin
date: Sat, 26 Oct 2024 19:55:26 GMT
referrer-policy: same-origin
server: WSGIServer/0.2 CPython/3.8.10
vary: Accept
x-content-type-options: nosniff
x-frame-options: DENY

```

Request duration

1627 ms

Responses

Code	Description
200	Job status
	Example Value Model
	<pre>ScrapingJob <- job_id* string title: Job id maxLength: 255 minLength: 1 status string title: Status maxLength: 50 minLength: 1 start_time string(date-time) title: Start time readOnly: true last_updated string(date-time) title: Last updated readOnly: true }</pre>
400	Job ID is required
404	Job not found

POST /apis/scraping-job/start_scraping/ apis_scraping-job_start_scraping ^

Parameters

Name	Description
data * required	

object
(body)

```
{
  "get_sample_links": false,
  "update_prev_links": false
}
```

[Cancel](#)

Parameter content type
application/json

[Execute](#) [Clear](#)

Responses Response content type application/json

Curl

```
curl -X 'POST' \
'http://0.0.0.0:4501/apis/scraping-job/start_scraping/' \
-H 'accept: application/json' \
-H 'Content-Type: application/json' \
-H 'Cookie: _CSRF_TOKEN=JySB0Qs5TeVtfZ9gCSt3inFk4nShYAV35jH3Pu37d6GbIJWLw4RxpYLr4W9B6b5' \
-d '{' \
"get_sample_links": false,
"update_prev_links": false
}'
```

Request URL
http://0.0.0.0:4501/apis/scraping-job/start_scraping/

Server response

Code	Details
202	<p>Response body</p> <pre>{ "status": "Scraping started", "job_id": "4e67d5d2-ca5e-44d6-93cf-60bf3e1815eb" }</pre> <p>Copy Download</p> <p>Response headers</p> <pre>allow: POST,OPTIONS content-length: 77 content-type: application/json cross-origin-opener-policy: same-origin date: Sat,26 Oct 2024 19:55:02 GMT referrer-policy: same-origin server: uvicorn/0.2 Python/3.8.10 vary: Accept x-content-type-options: nosniff x-frame-options: DENY</pre> <p>Request duration 19492 ms</p>
Responses	
Code	Description
202	Scraping job started

Example Value | Model

```
ScrapingJob {
  job_id*           string          title: Job id
                      maxLength: 255
                      minLength: 1
  status            string          title: Status
                      maxLength: 50
                      minLength: 1
  start_time        string($date-time) title: Start time
                      readOnly: true
  last_updated      string($date-time) title: Last updated
                      readOnly: true
}
```

400 Bad Request

auth

POST	/auth/login/	auth_login_create
POST	/auth/logout/	auth_logout_create
POST	/auth/register/	auth_register_create
POST	/auth/token/	auth_token_create
POST	/auth/token/refresh/	auth_token_refresh_create

Models

```
CollectedLinks {
  id             integer
  title: Id
  readOnly: true
  name*          string
  title: Name
  minLength: 1
  url*           string($uri)
  title: Url
  maxLength: 200
  minLength: 1
  folder         string
  title: Folder
  x-nullable: true
}
```

```
    last_updated      string($date-time)
    title: Last updated
    readOnly: true
  job_id           string
    title: Job id
    x-nullable: true
}
```

```
ExportedResults ∨ {
  id              integer
    title: ID
    readOnly: true
  collected_link* CollectedLinks ∨ {
    id          Id ∨ integer
      title: Id
      readOnly: true
    name*       Name ∨ string
      title: Name
      minLength: 1
    url*        Url ∨ string($uri)
      title: Url
      maxLength: 200
      minLength: 1
    folder       Folder ∨ string
      title: Folder
      x-nullable: true
    last_updated Last_updated ∨ string($date-time)
      title: Last updated
      readOnly: true
    job_id       Job id ∨ string
      title: Job id
      x-nullable: true
  }
  url*          string($uri)
    title: Url
    maxLength: 200
    minLength: 1
  result*       string
    title: Result
    minLength: 1
  taxonomy_level* string
    title: Taxonomy level
    x-nullable: true
  downloaded_file string
    title: Downloaded file
    x-nullable: true
  last_updated   string($date-time)
    title: Last updated
    readOnly: true
  job_id         string
    title: Job id
    x-nullable: true
}
```

```
ScrapingJob >
```

```
StartScraping >
```

```
UserRegistration ∨ {
  username*      string
    title: Username
    pattern: "([\\w.\\-]+)$"
    maxLength: 150
    minLength: 1
    required: true
    description: Required. 150 characters or fewer. Letters, digits and @./+/-/_ only.
  email          string($email)
    title: Email address
    maxLength: 254
  password*     string
    title: Password
    minLength: 1
}
```

```
TokenObtainPair >
```

```
TokenRefresh >
```

INVALID [-]