

API DOC for HtmlParserApplication with request/response and samples

1) POST /api/submit

Description :- This API is used to submit a url for scanning.

API Fields :-

API	Submitting the url for scanning
Method	POST
Request	/api/submit
Status	200

Body Fields :-

Field	Type	Mandatory	Description
urlAddr	String	Y	A url with http or https as a protocol

Success Response Fields :-

Field	Type	Description
id	Long	Id for the scanned URL
hostname	String	Hostname for the URL
scanStatus	String (SUCCESS, FAILED, ERROR)	Scan status of the URL
ipAddress	String	Ip address for the URL
title	String	Website title
submittedOn	String	Date of the time of submitting the scanning

Failure Response Fields :-

Field	Description
code	Code for the exception occurred
message	Error message

Sample Request and Response :-

Request :- <http://localhost:8080/api/submit>

Body :- {"urlAddr": "https://google.co.in"}

Success Response :-

```
{
  "id": 91,
  "hostname": "google.co.in",
  "scanStatus": "SUCCESS",
  "ipAddress": "172.217.27.195",
  "title": "Google",
  "submittedOn": "Tue Jun 19 21:32:33 IST 2018"
}
```

Failure Response :-

Body :- {"urlAddr": ".co.in"}

```
{
  "message": "Invalid URL format, Please enter URL with http[s]",
  "code": "INVALID_URL_FORMAT"
}
```

2) GET /api/list

Description :- This API is used get the list of last 10 (or given size) scanned urls

API Fields :-

API	Get list of Scanned URLs
Method	GET
Status	200
Request	/api/list

Request Parameters :-

Field	Type	Mandatory	Description
fromScan	Long (in millies)	N	Filter list on submitted date from
toScan	Long (in millies)	N	Filter list on submitted date to
size	Integer (Default 10)	N	Size of the list

Success Response Fields :-

Field	Type	Description
id	Long	Id for the scanned URL
hostname	String	Hostname for the URL
scanStatus	String (SUCCESS, FAILED, ERROR)	Scan status of the URL
ipAddress	String	Ip address for the URL
title	String	Website title
submittedOn	String	Date of the time of submitting the scanning

Failure Response Fields :-

Field	Description
code	Code for the exception occurred
message	Error message

Sample Request and Response :-

Request :- <http://localhost:8080/api/list?fromScan=&toScan=&size=>

Success Response :-

```
[
  {
    "id": 92,
    "hostname": "{\"urlAddr\":\"https://google.co.in\"}",
    "scanStatus": "FAILED",
    "ipAddress": "",
    "title": "",
    "submittedOn": "2018-06-19 21:42:40.0"
  },
  {
    "id": 91,
    "hostname": "google.co.in",
    "scanStatus": "SUCCESS",
    "ipAddress": "172.217.27.195",
    "title": "Google",
    "submittedOn": "2018-06-19 21:32:33.0"
  }
]
```

3) GET /api/scan/<scan_id>

Description :- This API is used get the details of a scanned urls

API Fields :-

API	Get Details of scanned URL
Method	GET
Status	200
Request	/api/scan/<scan_id>

Request Parameters :-

Field	Type	Mandatory	Description
scan_id	Long	Y	Scan Id for the scanned URL

Success Response Fields :-

Field	Type	Description
url	String	Hostname of the url
redirectionUrl	String	Redirection url if any
body	String	Body text scanned from the html
imgCount	Long	Total image count on the html
linksCount	Long	Total links count on the html
title	String	Website title
submittedOn	String	Date of the time of submitting the scanning

Failure Response Fields :-

Field	Description
-------	-------------

code	Code for the exception occurred
message	Error message

Sample Request and Response :-

Request :- <http://localhost:8080/api/scan/91>

Success Response :-

```
[
  {
    "url": "google.co.in",
    "redirectionUrl": "https://www.google.co.in/",
    "submittedOn": "2018-06-19 21:32:33.0",
    "title": "Google",
    "body": "Images Sign in Google Instant is unavailable Press Enter to search Learn more
Google Instant is off due to connection speed Press Enter to search Press Enter to search
Report inappropriate predictions window gbar Google offered in India Privacy Terms Settings
Search settings Advanced search History Search Help Send feedback Advertising Business
About q c k 56 q c k 64 q 56 for var g 63 56 a left c right a left b My Account Search Maps
YouTube Play News Gmail Contacts Drive Calendar Google Translate Photos More Docs
Books Blogger Hangouts Keep",
    "imgCount": 2,
    "linksCount": 69
  }
]
```

Failure Response :-

```
[
  {
    "message": "Scan Id not found",
    "code": "EXPECTATION_FAILED"
  }
]
```