

Data Extraction Tool - API Manual	
Version:	1.6.6
Version Date:	26/03/2024

Document history:

Revision	Date	Changes
1	26/10/2020	First version
1.1	05/11/2021	CMT review
1.2	09/11/2021	Revisions added
1.3	11/11/2021	CMT review
1.4	12/11/2021	Revisions added
1.4.1	27/06/2022	Review and comment
1.5.0	05/07/2022	Support page URL corrected. Open API URL and help page added.
1.6.0	17/01/2023	Revisions added
1.6.1	24/01/2024	Revisions added
1.6.2	29/01/2024	Revisions added
1.6.2.a	30/01/2024	CMT Review
1.6.3	31/01/2024	Revisions added
1.6.4	05/03/2024	Revisions added
1.6.5	26/03/2024	Fixed bug in deleteExtraction() and Unified extraction statuses
1.6.5	26/03/2024	Revision added

1. INTRODUCTION

The aim of this document is to explain the usage of the CORDIS Data Extraction Tool REST API (DET API hereafter). This application programming interface is intended for those users who want to extract CORDIS data without using the regular interface. By using the API you can automate extractions based on your needs and feed exported the data directly to your applications.

For sporadic or low volume exports please use the [regular user interface](#) which is available to any user registered in CORDIS that have been given the rights.

Although the API is completely free of use, the Publications Office of the EU reserves the right to suspend temporary or permanently the grants to those users that violate CORDIS website [terms and conditions](#). CORDIS sets and enforces limits on your use of all its APIs (e.g. limiting the number of API requests that you may make or the number of users you may serve), in our sole discretion. You agree to, and will not attempt to circumvent, such limitations. If you would like to use any API beyond these limits, you must obtain CORDIS express consent (and CORDIS may decline such request or condition acceptance on your agreement to additional terms and/or charges for that use).

2. BEFORE FIRST USE

The [API Services page](#) lists your API key that will be required for any use. If you think your key may have been compromised, you can always regenerate it from the same page clicking on the refresh icon.



Figure 1: DET API page


Action	Outcome
	Regenerates the API key and calculate a new one. A confirmation message appears to warn that if this action is triggered the user will lose all previous extractions.

Table 1: API page actions

3. CONSIDERATIONS

There are some considerations to take into account when using the API:

- Only 5 extractions can be associated with an API key at any one time. To remove old extractions, the API method "*deleteExtraction*" can be used see section:4.4.
- Extractions are automatically deleted 24 hours after completion.
- There is a limitation of 25 000 records for normal users and, extractions with over 10 000 records may take longer to complete due to the larger final file size. For larger data extractions, please consider using the CORDIS datasets available on data.europa.eu or our SPARQL interface instead, respectively:

<https://data.europa.eu/data/datasets?query=cordis&locale=en&publisher=Publications%20Office%20of%20the%20European%20Union&page=1&limit=25&categories=TECH&sort=modified%2Bdesc,%20relevance%2Bdesc,%20title.en%2Basc>
<https://cordis.europa.eu/datalab/sparql-endpoint/en>

4. PUBLIC API METHODS

The available methods have been detailed online in on Open API help page. This can be accessed directly, or through an interactive UI:

Open API: <https://cordis.europa.eu/dataextractions/api-docs>

Open API UI: <https://cordis.europa.eu/dataextractions/api-docs-ui>

The methods listed below are a subset of those currently available.

4.1. GET /api/dataextractions/getExtraction

This endpoint is used to request a new data extraction from CORDIS. Only one data extraction can be requested at the time. A second call to this method will cancel the previous one (a message will be left to indicate that in the response).

```
https://cordis.europa.eu/api/dataextractions/getExtraction
```

4.1.1. Input parameters

Parameter	Input	Description and samples
query	string Mandatory	Any valid query in CORDIS format. You can check samples of query syntax by executing a search and then switching to “Edit query” mode. Sample: contenttype='project' AND (/project/relations/associations/relatedTopic/programme/code =='MSCA-RISE-2015') To find out more about how to search CORDIS content, please visit https://cordis.europa.eu/about/search/en
outputFormat	string Mandatory Possible values: 'xml','xlsx','json','csv' Multiple values are allowed (comma separated) optional (default=xml)	<ul style="list-style-type: none"> • xml: individual XML files representing each a result record. • csv, xlsx and json: only if one (and only one) of these collections are set up on the query parameter: 'project','programme','pubsum','deliverable','publication','brief','news'
key	string Mandatory	Fill with the key obtained from the username profile https://cordis.europa.eu/about/api/[lang]

Table 2: getExtraction parameters

4.1.2. Sample call

Params:

Parameter	Input
query*	contenttype='project' AND (programme/code='H2020') AND 'rocket'
outputFormat	'xml','xlsx','json','csv','summary'
Key	123azerty456789

** queries can be experimented from the search feature in CORDIS web. After selecting the filters and searching you can copy the q parameter from the URL, like in the following sample:*

`https://cordis.europa.eu/search?q=contenttype%3D%27project%27%20AND%20(programme%2Fcode%3D%27H2020%27)%20AND%20%27rocket%27&p=1&num=10&srt=Relevance:decreasing`

Corresponding API call:

```
https://cordis.europa.eu/  
/api/dataextractions/getExtraction?query=contenttype%3D%27project%27%20AND%20(p  
rogramme%2Fcode%3D%27H2020%27)%20AND%20%27rocket%27&outputFormat=xml,xlsx,json,  
csv&key=123azerty456789
```

4.1.3. Sample responses

The responses will be sent in JSON format. It will contain the `taskID` of the extraction process. This `taskID` can be used later to retrieve the status of the running extraction.

```
{  
  "status": true,  
  "payload": {  
    "taskID": 64625437  
  }  
}
```

If a user had a previous extraction `taskID` running, it will be informed as “`cancelledExtraction`” along with the previous query. See sample below:

```
{  
  "status": true,  
  "payload": {  
    "taskID": 64625438,  
    "cancelledExtraction": {  
      "taskID": 64625437,  
      "query": "contenttype='project' AND (programme/code='H2020') AND 'rocket'"  
    }  
  }  
}
```

The extraction launched with this API method is scheduled and executed in background. Please take into account that large extractions can take hours to complete (approx. 4h per 25 000 records with all formats selected). When the extraction finishes, one of the following email templates are sent to user email (linked to your EU login account) depending on the result (completed or failed):

COMPLETED:

Subject: Your data extraction is ready (task ID: [taskId])

Dear [First Names] [Last Names],

The data extraction you requested on [EXTRACTION DATE] has finished.

You can download the result here:

[\[download link\]](#)

The files are also available in the [Data extractions section](#) of your CORDIS profile.

Kind regards,

The CORDIS Team

CORDIS is managed by the Publications Office of the European Union

FAILED:

Subject: Your data extraction has (task ID: [taskId])

Dear [First Names] [Last Names],

The data extraction you requested on [EXTRACTION DATE] failed due to technical issues.

The extraction can be relaunched in the [Data extractions section](#) of your CORDIS profile.

If you have more issues after relaunching the extraction, please contact our [Help Desk](#) and provide the Task ID of the extraction (you can find it in the subject of this email).

Kind regards,

The CORDIS Team

CORDIS is managed by the Publications Office of the European Union

4.2. GET /api/dataextractions/getExtractionStatus

This endpoint is used to find out the status and progress level of a given extraction. Once the extraction is finished an email will be sent with the URIs of the files to be downloaded.

4.2.1. Input parameters

Parameter	Input	Description and samples
taskId	integer Mandatory	ID of the process obtained from getExtraction response. Sample: 64625447
key	string Mandatory	Fill with the key obtained from the username profile https://cordis.europa.eu/about/api/[lang]

Table 3: getExtractionStatus parameters

4.2.2. Sample call

Params:

Parameter	Input
taskId	64625447
Key	123azerty456789

Generated call:

```
https://cordis.europa.eu/api/dataextractions/  
getExtractionStatus?taskId=64625447&key=123azerty456789
```

4.2.3. Sample Response

The response will be sent in JSON format. It will contain the status of the extraction. The progress field can be "Ongoing", "Finished" or "Failed". The response will also contain the taskID field, the query used for the extraction, the total number of records (numberOfRecords), the number of processed records (numberOfProcessedRecords) and the remaining time (remainingTime) and average speed (averageSpeed). Some of the fields may not be available or empty (e.g. destinationFileUri will be empty for all status except successfully "Finished"), depending on the progress of the extraction.

The following sample shows the progress of a running extraction:

```
{
  "status": true,
  "payload": {
    "taskID": 64625447,
    "destinationFileUri": "",
    "progress": "Ongoing",
    "query": "contenttype='project' AND (programme/code='H2020') AND 'biology'",
    "numberOfRecords": "3071",
    "numberOfProcessedRecords": "675",
    "remainingTime": "00:15:22",
    "averageSpeed": "0.4s/record"
  }
}
```

The following response is a sample of successfully finished extraction, which adds the destination file URIs and their respective sizes.

```
{
  "status": true,
  "payload": {
    "taskID": 64625447,
    "progress": "Finished",
    "query": "contenttype='project' AND (programme/code='H2020') AND 'biology'",
    "numberOfRecords": "3071",
    "numberOfProcessedRecords": "675",
    "remainingTime": "00:15:22",
    "averageSpeed": "0.4s/record",
    "destinationFileUri": [
      "\\export\\ec9db0f05165ec33a5ddcd2b8bb57e7efbb401c6\\64625448\\EXTRACTION_1673477926_20211027143328526.zip"
    ],
    "files": [
      "\\export\\ec9db0f05165ec33a5ddcd2b8bb57e7efbb401c6\\64625448\\csv.zip",
      "\\export\\ec9db0f05165ec33a5ddcd2b8bb57e7efbb401c6\\64625448\\json.zip",
      "\\export\\ec9db0f05165ec33a5ddcd2b8bb57e7efbb401c6\\64625448\\xlsx.zip",
      "\\export\\ec9db0f05165ec33a5ddcd2b8bb57e7efbb401c6\\64625448\\xml.zip"
    ],
    "filesSizes": [
      "(14.77M) ",
      "(16.50M) ",
      "(18.68M) ",
      "(63.35M) "
    ]
  }
}
```

}

For failed extractions the response is similar, but in this case the number of processed records will be less than the number of records and the “progress” value will be “Failed” instead of “Finished”. If this is the case it means that the process ended unexpectedly, and it is recommended to relaunch the extraction again with *getExtraction*.

4.3. GET /api/dataextractions/cancelExtraction

This endpoint is used to cancel an on-going extraction. As there can only be one extraction running per user key, this is the only parameter required for the method.

4.3.1. Input parameters

Parameter	Type	Description and samples
taskId	integer Mandatory	ID of the process obtained from getExtraction response. Sample: 64625447
key	string Mandatory	Fill with the key obtained from the username profile https://cordis.europa.eu/about/api/[lang]

Table 4: cancelExtraction parameters

4.3.2. Sample call

Params:

Parameter	Input
taskId	64625447
Key	123azerty456789

Generated call:

```
https://cordis.europa.eu/api/dataextractions/  
cancelExtraction?taskId=64625447&key=123azerty456789
```

4.3.3. Sample Response

The response will be sent in JSON format. It will return a message to the user saying the extraction was cancelled. A new extraction can be sent just after this is received.

Sample:

```
{  
  "status": true,  
  "payload": {  
    "taskId": 64625450,  
  }  
}
```



```

    "progress": "Cancelled",
    "query": "contenttype='project' AND (programme/code='H2020') AND 'biology'",
    "message": "Your extraction was cancelled correctly"
  }
}

```

4.4. DELETE /api/dataextractions/deleteExtraction

This endpoint is used to delete an extraction.

4.4.1. Input parameters

Parameter	Type	Description and samples
taskId	integer Mandatory	ID of the process obtained from getExtraction response. Sample: 64625447
key	string Mandatory	Fill with the key obtained from the username profile https://cordis.europa.eu/about/api/[lang]

Table 5: deleteExtraction parameters

4.4.2. Sample call

Params:

Parameter	Input
taskId	64625447
Key	123azerty456789

Generated call:

```

https://cordis.europa.eu/api/dataextractions/
deleteExtraction?taskId=64625447&key=123azerty456789

```

4.4.3. Sample Response

The response will be sent in JSON format. It will return a message to the user saying the extraction was deleted. A new extraction can be sent just after this is received.

Sample:

```

{
  "status": true,
  "payload": {
    "taskId": 64625450,
    "progress": "Deleted",
    "query": "contenttype='project' AND (programme/code='H2020') AND 'biology'",

```

```
    "message": "Your extraction was deleted correctly"  
  }  
}
```