

# Uploading document images | API

The path `/rest/uploads` includes endpoints that allow to import a document into Transkribus.

```
https://transkribus.eu/TrpServerTesting/rest/  
uploads?collId={collectionID}
```

A `POST` request to this endpoint creates a new upload process on the server. It is mandatory to set the query parameter `collId` which must include the ID of a collection where the user has write access.

If the header specifies `application/json` then a JSON object of the following form is expected:

```
{  
  "md": {  
    "title": "Bentham Box 35",  
    "author": "Jeremy Bentham",  
    "genre": "Notes",  
    "writer": "Secretary"  
  },  
  "pageList": {"pages": [  
    {  
      "fileName": "035_320_001.jpg",  
      "pageXmlName": "035_320_001.xml",  
      "pageNr": 1,  
      "imgChecksum": "9d531932c8e24d5a5  
dc13c92063698c9",  
      "pageXmlChecksum": "b644a9c34a65e  
e07c1c576194e720b4a"  
    },  
    {  
      "fileName": "035_321_001.jpg",  
      "pageXmlName": "035_321_001.xml",  
      "pageNr": 2,  
      "imgChecksum": "e3ae1a862b9cd53cc  
87c9325d2502547",  
      "pageXmlChecksum": "8ba4758b8b8d5"
```

```
df562e25809692be340"  
  
    }  
  
  ]}  
  
}
```

Besides some basic (optional) metadata, this object defines the structure of the document to upload including the filenames to expect.

A page object just has to have a fileName and a pageNr. All other fields are optional! The checksums must be computed with MD5, if used.

The response to this request will return an enriched object of the same type. It will include a unique upload ID

(field `uploadId` ) that is to be used for the following requests.

```
https://transkribus.eu/TrpServerTesting/rest/  
uploads/{uploadId}
```

This endpoint is used to `PUT` the files for each page to

Transkribus. Note, that the path now includes the `uploadId` from the response of the initial request.

The `Content-Type` of each request has to

be `multipart/form-data` and it must include the

complete data for one page, i.e. if a `pageXmlName` was set in the given structure object, then the image as well as the XML have to be delivered. It depends on the used library

whether the `Content-Type` has to be set explicitly. Please refer to the respective documentation on multipart requests. The body part names to be used

are `img` and `xml` respectively and both should be sent

as `application/octet-stream`.

If checksums have been defined, then the server will check the files upon each request and respond with `200` only if the transmission was flawless.

A `GET` request to this path can be used to check the status of the upload process intermediately. Once all files have been delivered successfully, the server will automatically start the ingest process. After the last `PUT` request is accepted, the returned object will include a field `jobId` that can be used to monitor the ingest process via `GET` requests to

```
https://transkribus.eu/TrpServerTesting/rest/  
jobs/{id}
```