

Hi all,

I am preparing a Talend job which will fetch data from web services using API. The structure of the web service is like this:

- Parent Folder (We have the URL for parent folder) Which contains Sub-Folders and inside Sub-Folders we have other Sub-Folders/ Documents.

Steps which we need to implement:

PS: I have changed the URLs (As it is confidential) (Given some random URLs for examples)

Step 1 : After hitting the Parent Folder URL using *tRestClient* Component, we will get a response in JSON format. Which includes information like:

- {
- "Name": "FOLDER_NAME"
- "CreateDate": "2017-02-10T15:24:44.87Z",
- "CreatedBy": "xyz@pqr.com",
- "UpdatedDate": "2019-01-18T19:34:26.723Z",
- "UpdatedBy": "shgqc@fqwefe.com",
- "Description": "",
- "ParentFolder": {
 "Href": "<https://gfqahfkcbak.com/5ag879a08-8220-e711-8a32-6c3be5a75f4f>"
- },
- "BrowseDocumentsUrl": "<https://dahgdiqu.com/8651/87caghbcas-acjhva97e-e711-9c0d-3ca82a1e3f41>",
 "AccessLevel": {
 "See": true,
 "Read": true,
 "Write": true,
 "Move": true,
 "Create": true,
 "SetAccess": true
- },
- "**Documents**": {
 "Href": "<https://gfqyegfe.com/assghvq/documents/98746afe-a97e-e711-9c0d-3ca82a1e3f41/documents>"
- },
- "**Folders**": {
 "Href": "<https://shdgfwbf.w.com/jyasfggj/folders/8f99d72e-a97e-e711-9c0d-3ca82a1e3f41/folders>"
- },
 "ShareLinks": {
 "Href": "some link"
- },
 "CreateDocumentHref": "some link",
 "Href": "some link".
- } + { "Name" : "SECOND_FOLDER_NAME" All details } + { "Name" : "THIRD_FOLDER_NAME" All details }

Step 2: From the above response we have to extract/take **Documents Href link** and **Folder Href Link** (Which will be the input for next *tRestClient* component.)

Step 3: With the **Documents Href link**: we will hit the API using *tRestClient* and in Response we get this :

- ```

{"Items":[
{"Name":"XYZ.docx",
"CreateDate":"2018-08-25T14:12:40.08Z",
"CreatedBy":"Workflow",
"UpdatedDate":"2019-07-15T14:12:40.08Z",
"UpdatedBy":"Workflow",
"Description":"TSomething",
"ParentFolder":{"Href":"https://dafduja.com/aaaff/folders/7264gvrbrgr67283--e911-9c25-3ca82a1e3f41"},
"HistoryItems":{"Href":"https://dqtgfugfq .com/qyfdqj/documents/23523jrgv2r-d4b2-e911-9c25-3c2a1e2213f3/historyitems"},
"AccessLevel":{"See":true,
"Read":true,
"Write":true,
"Move":true,
"Create":true,
"SetAccess":true},
"PageCount":2,
"Lock":{"Href":"https://ahfgtqyhqf.ajfgc /documents/32fwea8b10-d4b2-e911-9c25-3ca82a124t2tg4g1/lock"},
"PreviewUrl":{"Href":"https://ahfgtqyhqf.ajfgc /documents/32fwea8b10-d4b2-e911-9c25-3ca82a124t2tg4g1/lock"},
"Versions":{"Href":"https://ahfgtqyhqf.ajfgc /documents/32fwea8b10-d4b2-e911-9c25-3ca82a124t2tg4g1/lock"},
"ShareLinks":{"Href":"https://ahfgtqyhqf.ajfgc /documents/32fwea8b10-d4b2-e911-9c25-3ca82a124t2tg4g1/lock"},
"DocumentProcessTrackingActivities":{"Href":"https://ahfgtqyhqf.ajfgc /documents/32fwea8b10-d4b2-e911-9c25-3ca82a124t2tg4g1/lock"},
"DocumentReminders":{"Href":"https://ahfgtqyhqf.ajfgc /documents/32fwea8b10-d4b2-e911-9c25-3ca82a124t2tg4g1/lock"},
"RelatedDocuments":{"Href":"https://ahfgtqyhqf.ajfgc /documents/32fwea8b10-d4b2-e911-9c25-3ca82a124t2tg4g1/lock"},
"WorkItems":{"Href":"https://ahfgtqyhqf.ajfgc /documents/32fwea8b10-d4b2-e911-9c25-3ca82a124t2tg4g1/lock"},
"DownloadDocumentHref":{"Href":"https://ahfgtqyhqf.ajfgc /documents/32fwea8b10-d4b2-e911-9c25-3ca82a124t2tg4g1/lock"},
"NativeFileSize":845252,
"PdfFileSize":325156,{"Href":"https://ahfgtqyhqf.ajfgc /documents/32fwea8b10-d4b2-e911-9c25-3ca82a124t2tg4g1/lock"},} + {Other document details if available} + +
.....

```

**Step 4:** From above we have to take "**HistoryItems**":{"Href":"https://dqtgfugfq .com/qyfdqj/documents/23523jrgv2r-d4b2-e911-9c25-3c2a1e2213f3/historyitems"}, URL and hit it. Whatever it gives in Response we have to store that in DB. That is the final thing which ends one line of processing.

**Step 5:** From step 2, We will take Folder link and hit it. Then Again we will get Document + Folder Link. So we have to repeat From *Step 2 to Step 5*. Until we touch all documents present in the web service.

For better understanding I have prepared a Sketch of hierarchy in that web service: ( Which explains level of folders and documents inside it.) Please have a look on that once.

My need:

**Can anyone here please tell me the approach and Talend components which will help me to accomplish this task.**

**So I need a proper approach and list of Talend Components which can do this task.**

***Thank you in advance.***

***Note: I am new to this family and trying to learn and implement more.***

***Best Regards***

***Shahbaz***

Folder may contain Documents or subFolder. But at last there will be documents only on each line of processing