-Go to [Anaconda | Individual Edition](https://www.anaconda.com/products/individual-d)

-Download for windows or mac

-Open .exe from downloads

-Do the setup as is (don’t select or unselect anything)

-Click windows icon and Open Anaconda Powershell Prompt and type the following:

(Below commands are copy pasted from [UAL-RE/LD-Cool-P: Python tool to enable data curation (github.com)](https://github.com/UAL-RE/LD-Cool-P) with some changes)

-------------------------------------------------------------------------------------------------

$ conda create -n curation python=3.8.8

Click y when Proceed ([y]/n) option pops up

$ conda activate curation

(The above command creates a folder curation whose path is at C:\Users\username\Anaconda3\envs\curation)

With the activated conda environment, next clone the [UA Libraries' forked copy of figshare](https://github.com/UAL-RE/figshare) and install with the setup.py script:

Now, install git (if you haven’t already). Go to [Git - Downloads (git-scm.com)](https://git-scm.com/downloads) and click Download, open the .exe file from downloads and setup Git with all the default options.

(The path below will be C:/Users/username/Anaconda3/envs/curation, where username is what you have to enter, alternately you could go to the folder where curation folder was created, and copy paste the path below)

(curation) $ cd /path/to/parent/folder (this is the path to curation folder)

(curation) $ git clone https://github.com/UAL-RE/figshare.git

(The path below will be C:/Users/username/Anaconda3/envs/curation/figshare, where username is what you have to enter, alternately you could go to the folder where curation folder was created and copy paste the path below)

(curation) $ cd /path/to/parent/folder/figshare

(curation) $ (sudo) python setup.py develop

Then, clone this repository (LD-Cool-P) into the parent folder and install with the setup.py script:

(The path below will be C:/Users/username/Anaconda3/envs/curation/, where username is what you have to enter, alternately you could go to the folder where curation folder was created and copy paste the path below or type cd.. to go up one directory from the anaconda command prompt)

(curation) $ cd /path/to/parent/folder

(curation) $ git clone https://github.com/UAL-RE/LD-Cool-P.git

(The path below will be C:/Users/username/Anaconda3/envs/curation/LD-Cool-P, where username is what you have to enter, alternately you could go to the folder where LD-Cool-P folder was created, and copy paste the path below or type cd LD-Cool-P to go inside the curation directory from the anaconda command prompt)

(curation) $ cd /path/to/parent/folder/LD-Cool-P

(curation) $ (sudo) python setup.py develop

This will automatically install the required pandas, requests, numpy, jinja2, tabulate, and html2text packages.

You can confirm installation via conda list

(curation) $ conda list ldcoolp

You should see that the version is 1.1.6.

---------------------------------------------------------------------------------------------

-Move UPack\_v2.py (attached in this email) to the curation folder in ldcoolp: C:\Users\username\anaconda3\envs\curation\LD-Cool-P\ldcoolp\curation

-Now, go to start menu and open Spyder(Anaconda3) app

-Go to “File” and pick “New File”

-Copy the following code:

#----------------------------------------------------------------------------------------------------------------------------

import os

from os.path import exists

from figshare.figshare import Figshare

from ldcoolp.curation.UPack\_v2 import ObjFormatter

from ldcoolp.curation import retrieve

#Enter article id: this is the last number in the "cite" on data.lib.vt.edu

article\_id=15834846

# Enter Ingest Accession Number from the spreadsheet:

IngestAccessionNumber= "I00156"

#Enter Requestor name

Requestor="ThorneE"

#Enter corresponding author name

CorrespondingAuthor="ThorneE"

#Enter version number

Version="01"

#Enter date ingested in YYYYMMDD format

DateIngested= "20210823" #in YYYYMMDD format

#Enter your token

tn= '7f22a'

data\_directory1=f"{IngestAccessionNumber}"

data\_directory2=f"{IngestAccessionNumber}\_{Requestor}\_{CorrespondingAuthor}\_{Version}\_{DateIngested}"

data\_directory\_path=os.path.join(data\_directory1,data\_directory2)

fs=Figshare(token=tn,private=True)

z1=retrieve.download\_files(article\_id, fs, data\_directory=data\_directory\_path, metadata\_directory=data\_directory\_path)

u2='https://api.figshare.com/v2/account/articles/'+str(article\_id)

IngestMetadata=os.path.join(data\_directory\_path,f"{IngestAccessionNumber}\_IngestedMetadata"+'.gz')

z=retrieve.private\_file\_retrieve(url=u2, filename=IngestMetadata, token=tn)

#**Below addition made on Sept 22 2021**, it calls parts of modified UPACK\_v2 to bag and tar files

myobj=ObjFormatter#()

tarf=data\_directory1

za=ObjFormatter.run\_bagit(bagsdir=tarf)

zt=myobj.run\_tar(tarfolder=tarf)

#-----------------------------------------------------------------------------------------------------------------------

-Go to “File” and “Save as” and go to “C:\Users\username\anaconda3\envs\curation” or the path of the curation folder and save as “create\_ingest\_bag.py”

--Go to file and go to C:\Users\padma\anaconda3\envs\curation\LD-Cool-P\ldcoolp\curation and open retrieve.py

--Go to the last line (in my case this is at number 151 and put a hash before it like below:

# permissions.curation(dir\_path, mode=0o555) # read and execute only

The above changes the permissions on the .csv and .json files so that they can be bagged by UPACK

-Now, save the attached UPack\_v2.py to the curation folder at C:\Users\padma\anaconda3\envs\curation\LD-Cool-P\ldcoolp\curation in my case

-Now run the create\_ingest\_bag.py after filling in the article id, IngestAccessionNumber,Requestor, CorrespondingAuthor,Version,DateIngested, token and the ingest record will be created in the curation folder

If you get the following error

------------------------------------------

File "C:\Users\padma\Anaconda3\lib\site-packages\redata-0.4.1-py3.8.egg\redata\commons\logger.py", line 3, in <module>

from os import path, uname, chmod, mkdir

ImportError: cannot import name 'uname' from 'os' (C:\Users\padma\Anaconda3\lib\os.py)

-----------------------------------------------------

Then, click on “File "C:\Users\padma\Anaconda3\lib\site-packages\redata-0.4.1-py3.8.egg\redata\commons\logger.py", line 3, in <module>” and make the following changes:

On line 3 change

from os import path, uname, chmod, mkdir

to

#from os import path, uname, chmod, mkdir

from os import path, chmod, mkdir

from platform import uname

Save the File using save icon on the top or Ctrl+S or File->Save

Run create\_ingest\_bag.py again

The ingest record and metadata is now created at “C:/Users/username/Anaconda3/envs/curation” as I00156 for the filled article\_id

**Directions for publication folder:**

-Go to start menu and open Spyder(Anaconda3) app if its not already open

-Go to “File” and pick “New File”

-Copy the following code:

#-----------------------------------------------------------------------------------------------------------------------------------------

"""

Created on Tue Sep 28 09:41:18 2021

@author: padma

"""

import os

from os.path import exists

import json

from ldcoolp.curation import retrieve

from ldcoolp.curation import metadata

#Enter article id for published articles:this is the also last number in the "cite" on data.lib.vt.edu

article\_id=16628596

#Enter your token below

tn='7f22a'

#Enter published accession number from the spreadsheet

PublishedAccessionNumber= "P001555550"

#Enter requestor name

Requestor="BalantrapuN"

#Enter corresponding author name

CorrespondingAuthor="BalantrapuN"

#Enter version

Version="01"

#Enter published date in YYYYMMDD format

DatePublished= "20210922"

#Enter your token

tn='7f22a'

data\_directory1=f"{PublishedAccessionNumber}"

data\_directory2=f"{PublishedAccessionNumber}\_{Requestor}\_{CorrespondingAuthor}\_{Version}\_{DatePublished}"

data\_directory3=f"DisseminatedContent"

data\_directory\_path=os.path.join(data\_directory1, data\_directory2, data\_directory3)

u2='https://api.figshare.com/v2/articles/'+str(article\_id)

from figshare.figshare import Figshare

fs=Figshare(token=tn,private=False)

z1=retrieve.download\_files(article\_id, fs, data\_directory=data\_directory\_path, metadata\_directory=data\_directory\_path)

json\_out\_file1=f"{data\_directory\_path}/{PublishedAccessionNumber}\_PublishedMetadata.json"

json\_response1=fs.get\_article\_details(article\_id,version=None)

if not os.path.exists(json\_out\_file1):

with open(json\_out\_file1, 'w') as f:

json.dump(json\_response1,f,indent=4)

else:

print(f"File exists: {json\_out\_file1}")

if overwrite:

print("Overwriting!")

with open(json\_out\_file1, 'w') as f:

json.dump(json\_response1,f,indent=4)

#----------------------create VTCurationServicesActions

data\_directory4=f"VTCurationServicesActions"

data\_directory\_path2=os.path.join(data\_directory1,data\_directory2,data\_directory4)

os.mkdir(data\_directory\_path2)

print("Directory '% s' created" % data\_directory4)

#-----------------------------------------------------------------------------------------------------------------------------------------

-Go to “File” and “Save as” and go to “C:\Users\username\anaconda3\envs\curation” or the path of the curation folder and save the above code as “create\_publication\_bag.py”

-Fill the fields in the above code and run create\_publication\_bag.py