**Send images to DCC**

This document covers how to send images to DCC. This is a subprocess of sending data/procedure to DCC. Check out the document “To upload KOMP procedures to the DCC.doc” in <https://jacksonlaboratory.sharepoint.com/sites/ResearchIT/LIT%20Documents/Forms/AllItems.aspx?ga=1&id=%2Fsites%2FResearchIT%2FLIT%20Documents%2FApplications%2FKOMP&viewid=6e0da40e%2D1ce6%2D40cd%2D8314%2D4e68b3ed8e21> for more information. Before starting this process, make sure you have finished step 1 to 13. As of November 2023, all images should go to server [jlkinternal@bhjlk02lp.jax.org](mailto:jlkinternal@bhjlk02lp.jax.org).

Design

GitHub repository:

[*https://github.com/TheJacksonLaboratory/KOMP\_Images\_ETL*](https://github.com/TheJacksonLaboratory/KOMP_Images_ETL)

There are two main sources of images for KOMP team: 1) OMERO 2) Phenotype Drive (//jax.org/jax/phenotype). The filename format). Therefore, I separate the upload process into following parts:

1. *download\_from\_omero.py*: Download and upload images from OMERO.
2. *download\_from\_drive.py*: Download and upload images from phenotype drive.
3. *Missing\_images\_report.py*: Create a report for failed/missing images.

All of above steps require querying table ‘komp.imagefileuploadstatus’ table of production database (rslims.jax.org:3306). The query result will look like the following:

A screenshot of a computer screen

Description automatically generated

, column ‘SourceFileName’ is the place to download images; column ‘DestinationFileName’ is the location of images on sftp server. These are the two most important column we are going to use.

part 1) and part 2) follow the following algorithms:

1. Query the database to get location to download image(SourceFileName), destination of the image (DestinationFileName), testcode (taskkey).
2. Iterate through the row retrieved from the table.
3. For each row, download images to a local directory (in my case, I download them into folder C:\Program Files\KOMP\ImageDownload\pictures\IMPC\_XXX\_XXX), with filename in column “DestinationFileName” as final filename. For example, if column “SourceFileName” in a row is “example\_1.jpg” and column “DestinationFileName” is “123456\_example\_1.jpg”, your final filename should be the same as the one shown on column “DestinationFileName”.
4. Upload the downloaded file to sftp server.
5. Update image upload status of the table.
6. Delete the file from your local file system.
7. Repeat step 2) to 6) until there is no row left.

Finally, the missing image report script(missing\_images\_report.py) will query the table for all failed upload records in past 21 days and generate a .csv file using “taskKey” column in the table. The generated .csv file will be on phenotype drive, path is [\\jax.org\jax\phenotype\KOMP\missing\_images](file://jax.org/jax/phenotype/KOMP/missing_images). The filename format will be “Missing\_images\_Month\_dd\_yyyy.csv”.

Things worth noticing

1. If you want to modify the settings of the project, edit *utils.py* under the directory “src”.
2. Be sure to check the table and server manually to verify whether the images are successfully uploaded or not next morning (after you run the app).
3. Check folder [\\jax.org\jax\phenotype\KOMP\missing\_images](file://jax.org/jax/phenotype/KOMP/missing_images) for new missing images report, if there is, make sure you deliver it to lab technicians.

Current Status

The app is currently hosted and actively running on server bhwin0236.jax.org (location: C:\Program Files\KOMP\SoftwareUtilities\Images\_Import\_Scripts). To run it on the server, open your command prompt, cd to “C:\Program Files\KOMP\ImageDownload\KOMP\_Images\_ETL”, then type “submit”, which is .bat file that triggers *download\_from\_omero.py, download\_from\_drive.py*. To set it up somewhere else, please follow the steps in the README file.