How to scape new papers for OSPREE: June 2019

You'll need:

1. Excel or other program that makes .xls or .csv files

2. ImageJ download for free from here https://imagej.net/Welcome. You'll also need to add the Figure$_{C}alibration.class, whichwi$
Now here's what to do:

   1. Copy the excel file ospree_newpapers for git repo (ospree/data/ospree_2019update) and make your own extension, for example, Dan would write "ospree_2019update_dmb". This will be the spread sheet you enter your data into and then in the future, someone will merge all of our files together into the master data.

   2. Familiarize yourself with each tab:

      (a) **meta_general**: metadata for each sheet

      (b) **source**: list of the paper we are working with. Bibliographic information and notes on usefulness for our purposes. Note the "ToDo" column, which tells you which figure or table to focus on. You may find other figures are better, these were from our initial quick read. Also pay attention to datasetID column, which tells you how you should enter the identifying information for each paper.

      (c) **study**: Details on each experiment within each paper; possibly only one line for a paper, if only one experiment is relevant. This sheet is useful for our overview of what kind of experimental manipulations were done.

      (d) **data_simple**: Aggregated data on each experimental treatment combination.

      (e) **data_detailed**: Detailed data for the experiment, with all relevant information filled out.

      (f) Responses may be pre/post treatment, time, or other. Temporal responses, such as days to 50% budburst, are fairly common. An example of an other type of response would be percent budburst, again fairly common.

      (g) **scratch**: For temporary formatting and manipulating data scraped from ImageJ.

      (h) The two most important tabs to fill out are **study** and **data_detailed**. We can aggregate data down to get the data_simple version later.

   3. Read your paper and fill out the information in the "study" and "data_detailed tab"

   4. Find the designated figure or table as noted in the source tab of the worksheet. Also note the datasetID column in source; use this as you fill out data_detailed and study.

   5. Take a screen shot of the figure and import into ImageJ, following Tim's instructions from lab meeting with a video of screen shots from git/ospree/notes/howtoscrape/Data Scraping Tutorial.mp4. Use the scratch tab to get data into the right format, and then copy into data_detailed. Fill out the study tab as best as possible to describe the experimental treatments used in each study within each publication.

   6. A note on entering response times: If the respvar of a study is "daystobudburst" enter a 1 into the "response" column, and fill in the recorded time in the "response.time" column