The goal of this project is the creation of a workflow that uses the DMS system to initiate a partially automated analysis using MetaboliteDetector. To do this, we will need to:

1. Add a selection menu in the DMS system to select one of N libraries (5-10 hrs)
2. Add a selection window in the DMS system to select seed ions for the analysis (5-15 hrs)
3. Generate a MetaboliteDetector initialization file from information stored in the DMS system (5-10)
4. Modify the existing MetaboliteDetector by removing the existing GUI and replace it with text based input (20-60 hrs)
5. Output an MSL library file for known compounds and another for the unknown ones. (3-5 hrs)

TOTAL 38-100 hrs. Please see notes below for variation in time.

Steps 1 and 2 should not too difficult. With help from the DMS team, this should be a few hours of work. Step 3 is also probably not too difficult. It would require storing the MetaboliteDetector input information so it can be retrieved in the critical parts of the program. The solution to the problem is well known. Step 4 could be difficult. It will require "disconnecting" the GUI from the MetaboliteDetector algorithms and "connecting" the algorithms to the text file input from step 3. Step 5 is partially completed and should require very little work unless there are major changes to the format of the output.

Potential problems

- I do not foresee major problems in the first two steps. The code experts are here and should be able to help if there is a problem. The only potential problem is file retrieval and access.
- The third step could be difficult if the necessary data is difficult to find and its completeness is difficult to ensure. These conditions make it difficult to know when the task is completed.
- The fourth step has a lot of possible problems
    - The GUI could be intertwined with the algorithms. This will make the separation of the two components difficult and error prone.
    - Identifying all of the GUI components may be time consuming.
    - Identifying "hidden" dependencies can be time consuming
    - Finding and replacing counterproductive code
- The fifth step could have some complications if the "known" output files require more information than the "unknown" files. This is a matter of finding where the information is located.