1. Simplify data input.

I have ~10 ROI for each cell (image) and up to 30 cells for one experiment. It is time consuming to run the find\_colclus.py script for all the excel sheets 1 by 1.

Solutions:

1. We can put all excel sheets in a subfolder and execute the find\_colclus.py to all. **Or**
2. ~~We can add column titles in a note file. For example: Q: DOC; R: Lr etc.~~
3. Sort out the excel sheets.

The clus-doc file should generate a subfolder, which includes excel sheets for all ROIs of one image file (such as 1.txt). “ROI\_1\_in\_cluster\_Ch1.xls”; “ROI\_1\_non\_cluster\_Ch1.xls”, “ROI\_2\_non\_cluster\_Ch1.xls”, “ROI\_2\_In\_cluster\_Ch1.xls”……   
(I think it is better to combine ROI\_1\_in\_cluster\_Ch1.xls”; “ROI\_1\_non\_cluster\_Ch1.xls”, to ROI\_all\_events\_Ch1.xls). The folder should structure like this Text

Description automatically generated

The user defines a date folder ( 20221215), and run number (such as 1, 2, 3 ,4 ), The script will go to the subfolder to retrieve the excel sheets and calculate the parameter for each ROI for each image file.

1. Calculation from the excel sheets.

Goal

Analyze the distribution of all events in noncluster, colocalized clusters, and non colocalized cluster.

The output in a excel sheet should contain the following information **per image ( or per 1.txt file)**.

~~Cell # (same as the #.txt file)~~; ROI number; number of total events; number of **nonclustered** events; number of events in colocalized clusters; number of events in nonlocalized clusters. Number of colocalized events; number of **nonclustered** colocalized events; number of colocalized events in colocalized clusters; number of colocalized events in non-colocalized clusters.

To calculate: For each ROI, each channel, two excel sheets are available:

ROI\_1\_in\_cluster\_Ch1.xls; ROI\_1\_non\_cluster\_Ch1.xls.

**It may be easier if the two excel sheets can be combined. Can be called ROI\_1\_all\_events\_Ch1.xls**

* Number of total events: total row number of “the two excels **or** the combined excel”.
* Number of non clustered events: Total row number of “ROI\_1\_non\_cluster\_Ch1.xls”; **or** Total row number of “V = 0 (cluster ID) in ROI\_1\_all\_events.xls”.
* Number of events in colocalized clusters: Total row number “w (Col) = 1 in ROI\_1\_in\_cluster\_Ch1.xls” **or** “Total row number w = 1 and V (cluster ID) ǂ 0 in ROI\_1\_all\_events.xls”
* Number of events in nonlocalized clusters: Total row number “w (Col) = 0 in ROI\_1\_in\_cluster\_Ch1.xls” **or** Total row number “w = 0 and V (cluster ID) ǂ 0 in ROI\_1\_all\_events.xls”
* Number of colocalized events: total row number “Q (DoC) > = 0.4 in both ROI\_1\_in\_cluster\_Ch1.xls and ROI\_1\_non\_cluster\_Ch1.xls” **or “**Q (DoC) > = 0.4 in ROI\_1\_all\_events.xls”.
* Number of **nonclustered** colocalized events: total row number “Q (DoC) > = 0.4 in ROI\_1\_non\_cluster\_Ch1.xls” **or** total row number “Q (DoC) > = 0.4 and V(ClusterID) = 0 in ROI\_1\_all\_events.xls”.
* Number of colocalized events in colocalized clusters: total row number “Q (DoC) > = 0.4 and W (col) = 1 in ROI\_1\_In\_cluster\_Ch1.xls” **or** total row number “Q (DoC) > = 0.4 and V(ClusterID) != 0 and W (col) = 1 in ROI\_1\_all\_events.xls”.
* Number of colocalized events in non-colocalized clusters: total row number “Q (DoC) > = 0.4 and W (col) = 0 in ROI\_1\_In\_cluster\_Ch1.xls” **or** total row number “Q (DoC) > = 0.4 and V(ClusterID) != 0 and W (col) = 0 in ROI\_1\_all\_events.xls”.