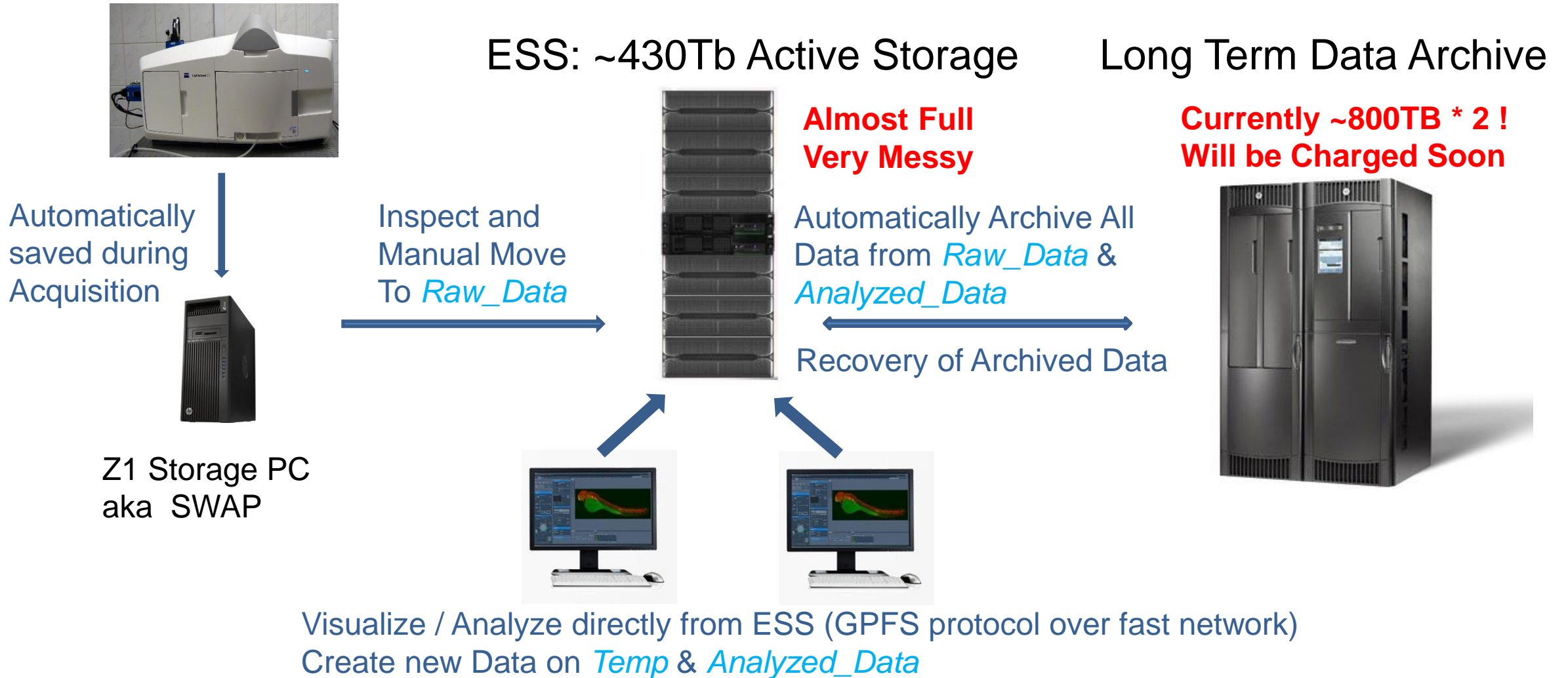

Cell Observatory Users Meeting

ESS Guidelines and Immediate Required Action Items

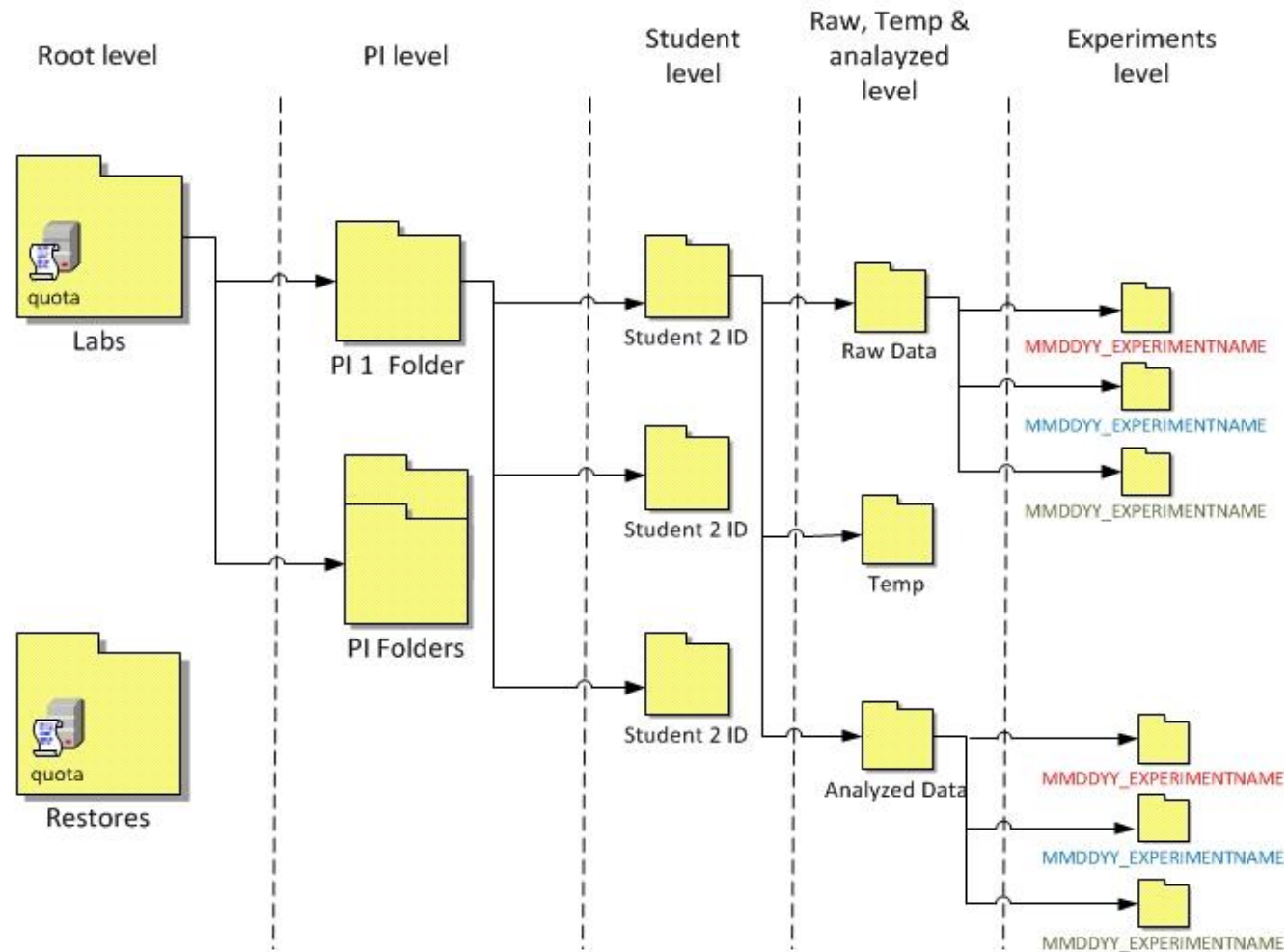
Aims

- Cell Observatory Data Storage Overview
- ESS Usage Guidelines – Active from tomorrow
- Required Immediate Action
 - 1st step By February 8th
 - 2nd step by February 26th

Cell Observatory Data Storage Overview

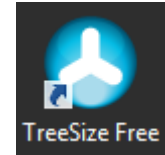


ESS Folder Structure



My ESS Disk Usage

Use TreeSize Free to inspect your ESS disk Usage



TreeSize Free - G:\Labs\BennyShilo\netast\ on [gpfs0]

File Scan View Expand Options Help TreeSize Professional

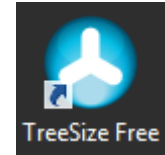
KB MB GB

Name	Size	Allocated	Files	Folders	% of Paren...	Last Change
100.0 % G:\Labs\BennyShilo\netast\ on [gpfs0]	60.6 TB	60.6 TB	16,629	263	100.0 %	1/2/2017
88.5 % Raw_Data	53.6 TB	53.6 TB	9,232	184	88.5 %	12/22/2016
11.4 % Analyzed_Data	6.9 TB	6.9 TB	1,114	76	11.4 %	1/2/2017
0.1 % Temp	43.4 GB	46.0 GB	6,283	0	0.1 %	11/11/2016

Free Space: 84.6 TB (of 430 TB) | 16,629 Files | 524288 Bytes per Cluster (GPFS)

My ESS Disk Usage

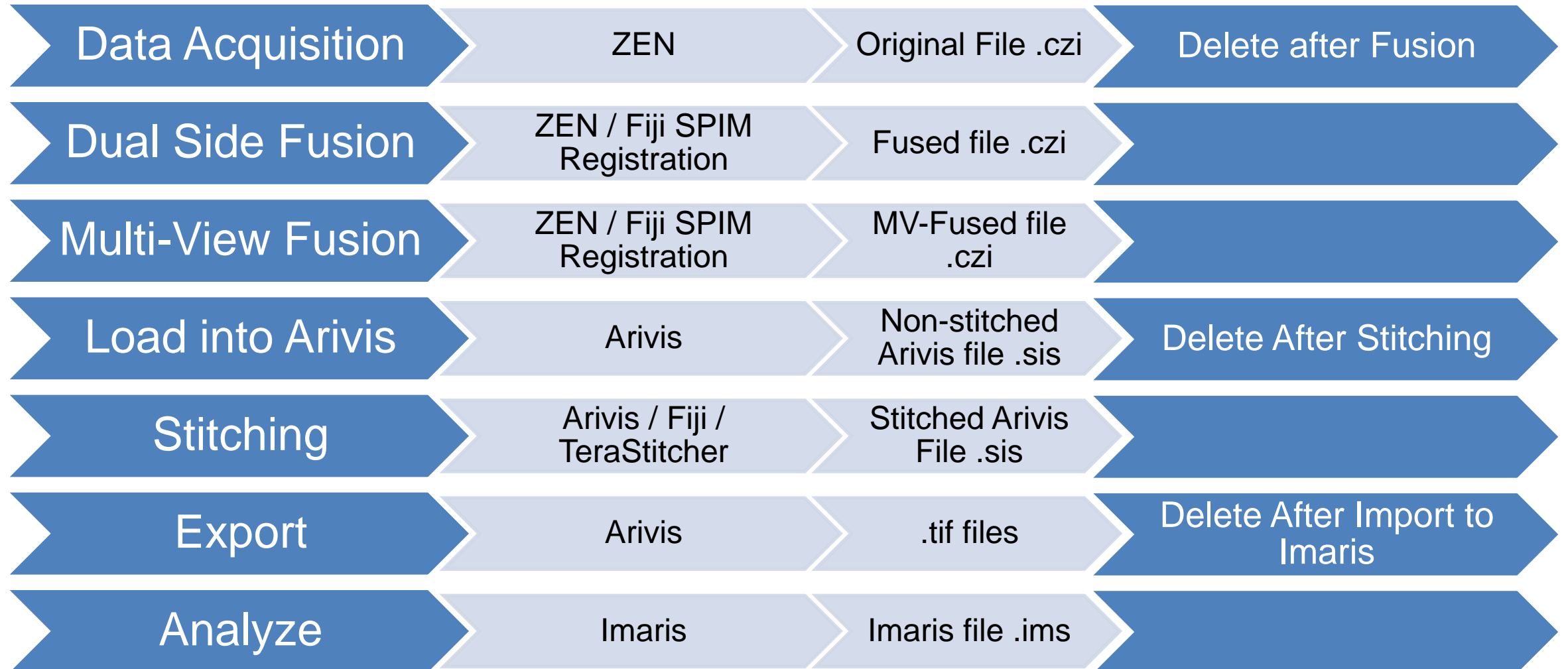
Use TreeSize Free to inspect your ESS disk Usage



Name	Size	Allocated	Files	Folders	% of Paren...	Last Change
100.0 % G:\Labs\BennyShilo\netast\ on [gpfs0]	60.6 TB	60.6 TB	16,629	263	100.0 %	1/2/2017
88.5 % Raw_Data	53.6 TB	53.6 TB	9,232	184	88.5 %	12/22/2016
38.4 % All_DATA_Before_G0	20.6 TB	20.6 TB	2,115	104	38.4 %	10/5/2016
22.8 % DIGFP	12.2 TB	12.2 TB	4,954	37	22.8 %	12/22/2016
17.6 % sna movies	9.4 TB	9.4 TB	605	20	17.6 %	7/19/2016
13.5 % snaMS2	7.3 TB	7.3 TB	778	4	13.5 %	12/18/2016
7.4 % wntDms2	4.0 TB	4.0 TB	764	8	7.4 %	8/21/2016
0.1 % Data for Paul	76.1 GB	76.1 GB	14	3	0.1 %	7/20/2016
0.1 % [1 Files]	28.8 GB	28.8 GB	1	0	0.1 %	6/16/2016
0.0 % UIR	5.8 MB	6.0 MB	1	1	0.0 %	11/26/2015
11.4 % Analyzed_Data	6.9 TB	6.9 TB	1,114	76	11.4 %	1/2/2017
49.8 % DI GFP	3.5 TB	3.5 TB	290	33	49.8 %	1/2/2017
32.7 % Data for Inna	2.3 TB	2.3 TB	728	28	32.7 %	11/6/2016
10.1 % snaMS2	714.8 GB	714.8 GB	5	2	10.1 %	11/20/2016
5.1 % wntDms2	365.2 GB	365.3 GB	44	3	5.1 %	11/20/2016
1.2 % 180416_sna_ms2	87.7 GB	87.7 GB	1	0	1.2 %	5/4/2016
0.8 % 220216	59.1 GB	59.1 GB	10	0	0.8 %	8/4/2016
0.3 % 200116_snaMS2_OfraTest	20.3 GB	20.3 GB	24	1	0.3 %	11/7/2016
0.0 % 030115_wntD_fish fixed with beads	2.2 GB	2.2 GB	5	0	0.0 %	10/25/2016
0.0 % ppts_and_courses	41.2 MB	43.5 MB	7	0	0.0 %	12/11/2016

Free Space: 84.6 TB (of 430 TB) | 16,629 Files | 524288 Bytes per Cluster (GPFS)

Typical Workflow: Data Duplication Problem



ESS Usage Guidelines

- Follow File and folder naming guidelines (meaningful name, no spaces, no special characters, don't save files directly under Raw/Analyzed/Temp)
- Raw Data should be always archived
- Never Duplicate Raw Data files on ESS (Under *Raw_Data* folder, or different user)
- Minimize (Saving of) Data Duplicates
- Dual Side Fusion and Multi-View Fusion (using ZEN) will be saved under *Raw_Data*
All Further Data Visualization & Analysis should be done in the user's *Temp* folder
- Carefully Select which Analyzed data to archive, **move** it to *Analyzed_Data* (**Don't copy**) – Charging will be based on size of archived data
- Use Matched Folders Names in *Raw_Data* / *Analyzed_Data* / *Temp*
- Quota / Time policy will be imposed

Active for new data from Jan 24th

Data Archive & Retrieval

- Data will not be saved on ESS forever
- All Data is stored on 2 copies
- Raw and Processed Archived Data can be recovered from the Tapes
- Note that any change to a file create another archived copy (analyzed data) Archived data will be charged by TB
- In order to retrieve files from the Tapes a full path should be provided
- To enable data retrieval a database application will be created
- Retrieval is into *Restore* folder

Full Data Retrieval Instructions will be posted soon

Data Annotation

- Storing the information about the datasets names and content is required to enable data retrieval
- A database application will help you keep track of your files and experiments.
- Impossible to save information at the single file level (optionally multiple files for experiment)
- **Path and description will be saved at the folder level only and not at the single file level**
- A new item will be added for each subfolder that contain data files
- Automatically saved info: Folder name, Full path, Folder creation date, Number of files, size, Data ownership
- The user is required to fill in additional information:
 - Experiment information (textual): conditions, date , number of files of each experiment
 - Anything that you think you or your PI will need 2 years from now to identify the experiment
 - Base File names, multiple files separated by comma

Full Data Annotation Instructions will be posted soon

Nested Folders, Single File per Experiment

The screenshot shows the TreeSize Free interface. The main window displays a directory tree for 'G:\ on [gpfs0] (Scan of 1/10/2017)'. The tree is as follows:

- 100.0 % stoler (150.8 TB, 150.8 TB, 14,544 files, 340 folders, 100.0% of parent, 1/10/2017)
 - 62.8 % Raw_Data (94.7 TB, 94.7 TB, 465 files, 114 folders, 62.8% of parent, 1/10/2017)
 - 98.7 % Inflamed (93.5 TB, 93.5 TB, 444 files, 104 folders, 98.7% of parent, 1/10/2017)
 - 12.0 % T-B_T OTII DsRed_B1 8High GFP (11.2 TB, 11.2 TB, 29 files, 4 folders, 12.0% of parent, 7/9/2016)
 - 38.8 % Day 8 (4.3 TB, 4.3 TB, 8 files, 0 folders, 38.8% of parent, 7/9/2016)
 - 32.7 % Day 5 (3.7 TB, 3.7 TB, 8 files, 0 folders, 32.7% of parent, 7/9/2016)
 - 18.1 % Day 3 (2.0 TB, 2.0 TB, 6 files, 0 folders, 18.1% of parent, 7/6/2016)
 - 22.8 % T-B_day3_mouse1_LN2_Inflamed.czi (472.4 GB, 472.4 GB, 1 file, 0 folders, 22.8% of parent, 7/6/2016)
 - 22.6 % T-B_day3_mouse1_LN1_Inflamed.czi (468.2 GB, 468.2 GB, 1 file, 0 folders, 22.6% of parent, 7/3/2016)
 - 21.3 % T-B_day3_mouse2_LN1_Inflamed.czi (440.2 GB, 440.2 GB, 1 file, 0 folders, 21.3% of parent, 7/3/2016)
 - 11.4 % T-B_day3_mouse1_LN2_Inflamed_DualSideFusion.czi (236.2 GB, 236.2 GB, 1 file, 0 folders, 11.4% of parent, 7/6/2016)
 - 11.3 % T-B_day3_mouse1_LN1_Inflamed_DualSideFusion.czi (234.1 GB, 234.1 GB, 1 file, 0 folders, 11.3% of parent, 7/4/2016)
 - 10.6 % T-B_day3_mouse2_LN1_Inflamed_DualSideFusion.czi (220.1 GB, 220.1 GB, 1 file, 0 folders, 10.6% of parent, 7/5/2016)
 - 10.5 % Day 0 (1.9 TB, 1.9 TB, 7 files, 0 folders, 2.0% of parent, 6/30/2016)
 - 11.6 % MD4 (234.1 GB, 234.1 GB, 1 file, 0 folders, 11.3% of parent, 7/4/2016)
 - 9.7 % YFP_AID_NP_TOMATO (234.1 GB, 234.1 GB, 1 file, 0 folders, 11.3% of parent, 7/4/2016)
 - 9.6 % Confetti_CGG_12days (234.1 GB, 234.1 GB, 1 file, 0 folders, 11.3% of parent, 7/4/2016)
 - 8.6 % EDU (234.1 GB, 234.1 GB, 1 file, 0 folders, 11.3% of parent, 7/4/2016)
 - 5.7 % PA B1_8 high NP_Tomato (234.1 GB, 234.1 GB, 1 file, 0 folders, 11.3% of parent, 7/4/2016)
 - 4.4 % Lyve1_12h_Blimp_YFP (234.1 GB, 234.1 GB, 1 file, 0 folders, 11.3% of parent, 7/4/2016)
 - 4.1 % Confetti_NP-OVA_12days (234.1 GB, 234.1 GB, 1 file, 0 folders, 11.3% of parent, 7/4/2016)
 - 3.6 % GFP_AID_Inflamed (234.1 GB, 234.1 GB, 1 file, 0 folders, 11.3% of parent, 7/4/2016)
 - 3.2 % CHIMERA50-50 (234.1 GB, 234.1 GB, 1 file, 0 folders, 11.3% of parent, 7/4/2016)
 - 3.0 % AID_Cre_Blimp-real mouse- (234.1 GB, 234.1 GB, 1 file, 0 folders, 11.3% of parent, 7/4/2016)
 - 2.8 % CD4blocking_AIDtomato (234.1 GB, 234.1 GB, 1 file, 0 folders, 11.3% of parent, 7/4/2016)
 - 2.6 % Blimp_flox_AID_Tomato (234.1 GB, 234.1 GB, 1 file, 0 folders, 11.3% of parent, 7/4/2016)
 - 2.6 % anti_DEC_OVA_AID_Tomato (234.1 GB, 234.1 GB, 1 file, 0 folders, 11.3% of parent, 7/4/2016)
 - 2.5 % Tomato_AID_Inflamed (234.1 GB, 234.1 GB, 1 file, 0 folders, 11.3% of parent, 7/4/2016)
 - 2.2 % anti TCR 14h (234.1 GB, 234.1 GB, 1 file, 0 folders, 11.3% of parent, 7/4/2016)
 - 2.0 % Bortezamib (234.1 GB, 234.1 GB, 1 file, 0 folders, 11.3% of parent, 7/4/2016)

The detailed file information pane for 'T-B_day3_mouse1_LN1_Inflamed_DualSideFusion.czi' shows:

 - Name: T-B_day3_mouse1_LN1_Inflamed_DualSideFusion.czi
 - Full Path: G:\Labs\ZivShulman\stoler\Raw_Data\Inflamed\T-B_T OTII DsRed_B1 8High GFP\Day 3\T-B_day3_mouse1_LN1_Inflamed_DualSideFusion.czi
 - Size: 234.1 GB
 - Allocated: 234.1 GB
 - % of Parent: 11.3%
 - Optical media size: 234.1 GB
 - Files: 1 Files
 - Folders: 0
 - Last Change: 7/4/2016
 - Last Access: 7/6/2016
 - Creation Date: 7/4/2016
 - Owner: stoler
 - Attributes: A
 - Compr.: 0.0%
 - Permissions: Administrators: full, BS_AdminGrp: full, BS_ZivShulman_Grp: +r+w+x, Domain Admins: full
 - Dir Level: 8
 - Avg. File Size: 234.1 GB
 - Path Length: 128

Free Space: 87.1 TB (of 430 TB) | 1 Files | 524288 Bytes per Cluster (GPFS)

Nested Folders, Multiple Files per Experiment

TreeSize Free - G:\ on [gpfs0]

File Scan View Expand Options Help TreeSize Professional

Name	Size	Allocated	Files	Folders	% of Paren...	Last Change
2.1 TB EliArama	2.1 TB	2.1 TB	420	33	0.6 %	10/01/2017
1.5 TB AvrahamLevy	1.5 TB	1.5 TB	19,307	41	0.5 %	15/01/2017
1.5 TB AlonChen	1.5 TB	1.5 TB	8,432	31	0.5 %	05/01/2017
1.5 TB aramot	1.5 TB	1.5 TB	8,432	30	100.0 %	05/01/2017
1.2 TB Raw_Data	1.2 TB	1.2 TB	6,441	23	82.5 %	05/01/2017
296.0 GB 140416CRFcrexAi9	296.0 GB	296.0 GB	2	1	23.5 %	17/04/2016
177.1 GB 08022016	177.1 GB	177.1 GB	4	0	14.0 %	11/02/2016
140.4 GB 26052016	140.4 GB	140.4 GB	5	0	11.1 %	29/05/2016
111.0 GB CAPTURE project	108.4 GB	111.0 GB	6,328	7	8.8 %	05/01/2017
45.6 GB 161221_brain2_z_tile2-12-16-31	45.6 GB	45.6 GB	235	1	41.1 %	23/12/2016
40.3 GB 161221_brain2_z_tile3x0_63_thickslice_12-44-23	38.7 GB	40.3 GB	3,748	0	36.3 %	21/12/2016
21.0 GB 27122016	20.2 GB	21.0 GB	1,960	1	18.9 %	05/01/2017
21.0 GB 161227_1b_3_hemisphere_Ai9_10-22-47	20.2 GB	21.0 GB	1,960	0	100.0 %	05/01/2017
11.5 MB 10-22-47_1b_3_hemisphere_Ai9_Ultrall[00x00]_C00_xyz-Table Z0000.ome.tif	11.0 MB	11.5 MB	1	0	0.1 %	27/12/2016
11.5 MB 10-22-47_1b_3_hemisphere_Ai9_Ultrall[01x00]_C00_xyz-Table Z0000.ome.tif	11.0 MB	11.5 MB	1	0	0.1 %	27/12/2016
11.0 MB 10-22-47_1b_3_hemisphere_Ai9_Ultrall[00x00]_C00_xyz-Table Z0001.ome.tif	10.6 MB	11.0 MB	1	0	0.1 %	27/12/2016
11.0 MB 10-22-47_1b_3_hemisphere_Ai9_Ultrall[00x00]_C00_xyz-Table Z0002.ome.tif	10.6 MB	11.0 MB	1	0	0.1 %	27/12/2016
11.0 MB 10-22-47_1b_3_hemisphere_Ai9_Ultrall[00x00]_C00_xyz-Table Z0003.ome.tif	10.6 MB	11.0 MB	1	0	0.1 %	27/12/2016
11.0 MB 10-22-47_1b_3_hemisphere_Ai9_Ultrall[00x00]_C00_xyz-Table Z0004.ome.tif	10.6 MB	11.0 MB	1	0	0.1 %	27/12/2016
11.0 MB 10-22-47_1b_3_hemisphere_Ai9_Ultrall[00x00]_C00_xyz-Table Z0005.ome.tif	10.6 MB	11.0 MB	1	0	0.1 %	27/12/2016
11.0 MB 10-22-47_1b_3_hemisphere_Ai9_Ultrall[00x00]_C00_xyz-Table Z0006.ome.tif	10.6 MB	11.0 MB	1	0	0.1 %	27/12/2016
11.0 MB 10-22-47_1b_3_hemisphere_Ai9_Ultrall[00x00]_C00_xyz-Table Z0007.ome.tif	10.6 MB	11.0 MB	1	0	0.1 %	27/12/2016
11.0 MB 10-22-47_1b_3_hemisphere_Ai9_Ultrall[00x00]_C00_xyz-Table Z0008.ome.tif	10.6 MB	11.0 MB	1	0	0.1 %	27/12/2016
11.0 MB 10-22-47_1b_3_hemisphere_Ai9_Ultrall[00x00]_C00_xyz-Table Z0009.ome.tif	10.6 MB	11.0 MB	1	0	0.1 %	27/12/2016
11.0 MB 10-22-47_1b_3_hemisphere_Ai9_Ultrall[00x00]_C00_xyz-Table Z0010.ome.tif	10.6 MB	11.0 MB	1	0	0.1 %	27/12/2016
11.0 MB 10-22-47_1b_3_hemisphere_Ai9_Ultrall[00x00]_C00_xyz-Table Z0011.ome.tif	10.6 MB	11.0 MB	1	0	0.1 %	27/12/2016
11.0 MB 10-22-47_1b_3_hemisphere_Ai9_Ultrall[00x00]_C00_xyz-Table Z0012.ome.tif	10.6 MB	11.0 MB	1	0	0.1 %	27/12/2016
11.0 MB 10-22-47_1b_3_hemisphere_Ai9_Ultrall[00x00]_C00_xyz-Table Z0013.ome.tif	10.6 MB	11.0 MB	1	0	0.1 %	27/12/2016
11.0 MB 10-22-47_1b_3_hemisphere_Ai9_Ultrall[00x00]_C00_xyz-Table Z0014.ome.tif	10.6 MB	11.0 MB	1	0	0.1 %	27/12/2016
11.0 MB 10-22-47_1b_3_hemisphere_Ai9_Ultrall[00x00]_C00_xyz-Table Z0015.ome.tif	10.6 MB	11.0 MB	1	0	0.1 %	27/12/2016
11.0 MB 10-22-47_1b_3_hemisphere_Ai9_Ultrall[00x00]_C00_xyz-Table Z0016.ome.tif	10.6 MB	11.0 MB	1	0	0.1 %	27/12/2016

Free Space: 106 TB (of 430 TB) | 5 Files | 524288 Bytes per Cluster (GPFS)

ESS Reorganization & Cleanup

Required Action Step 1: **by February 8**

- Delete all data that is
 1. Not needed (bad experiment, temporary)
 2. No need to archive it
- Reorganize the data

Create updated fresh archive

Required Action Step 2: **TBA, tentatively by February 26**

- Annotate data to enable retrieval
- Delete data that is not needed currently (once archived)

Required Action Step 1: **By February 8**

Delete unsuccessful experiments

- From all folders

Remove Duplicated Data Step1

- For all files stitched in Arivis – Delete the non-stitched Arivis files
- Delete all temporary tif files (exported from Arivis for import in Imaris)
- Remove all other un-needed data (eg. Duplicated files) from ESS and from WS

Reorganize Data

- Match folders names under *Raw_Data* / *Analyzed_Data* / *Temp*
- **Move** analyzed data that you do not want to archive into *Temp* folder

Required Action Step 2: TBA **By February 26**

Annotate Data to enable Data Retrieval

- We are currently building an application for this, more instructions will be posted soon

Remove Duplicated Data step 2

- For all files for which Dual-Side Fusion was done successfully – Delete the original files (before fusion)
- Remove all other un-needed data (eg. Duplicated files)

Remote Data Access and Further Info

Connecting to the ESS from your lab

- Open my computer → Map network drive → Folder → \\analysis80t.wismain.weizmann.ac.il\Labs\lab_name
- At security prompt enter:
 - username - wismain\userID
 - password - your WIS regular password

Further info can be found on the **Unit WIKI page**:

<https://bbcunit.atlassian.net/wiki/display/BIMGP/Advanced+Light+Microscopy+Unit>