

Oxford Update

Hierarchical Data, Tagging and Search

Douglas Russell

douglas.russell@bioch.ox.ac.uk

June, 2013



DEPARTMENT OF BIOCHEMISTRY
UNIVERSITY OF OXFORD



wellcometrust
Strategic Award

Micron
OXFORD

Table of contents

Update

- Processing Services
- Tagging Prototype
- C++ Gateway

HierarchialData

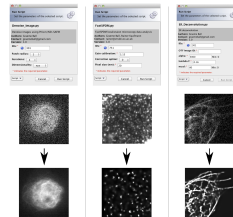
- Filesystem Hierarchy
- Important Information
- Tag Navigation
- Nested Common Search Criteria
- In Practice

Update

- ▶ Installation of large scale OMERO in production
- ▶ Proliferation and promotion of use in Davis lab and beyond
- ▶ Feedback to Dundee
- ▶ Processing Services
- ▶ Tagging Prototype
- ▶ C++ Gateway

Processing Services

- ▶ Denoising
- ▶ Deconvolution
- ▶ Localisation Microscopy
- ▶ Matalab Analysis



Tagging Prototype

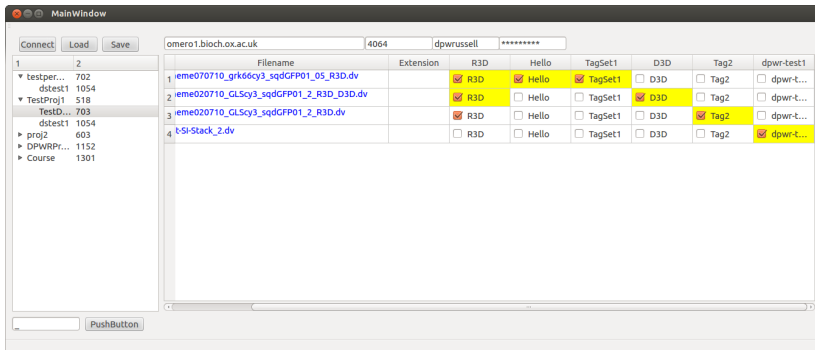


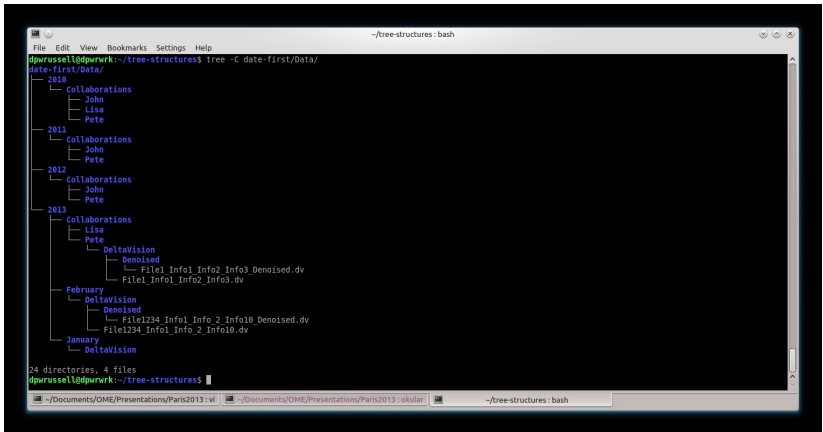
Figure: Prototype Demonstrating Auto-Tagging

C++ Gateway

- ▶ Friendlier C++ interface wrapping for OMERO Clients
- ▶ Based on and mirroring the Python Gateway

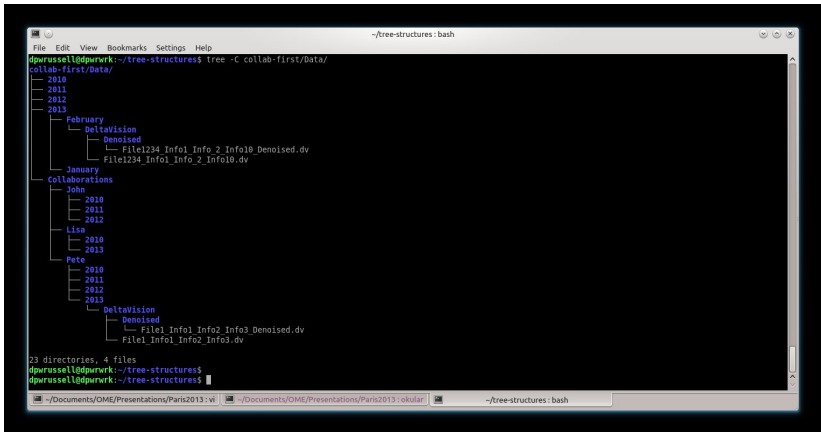
Hierarchical Data, Tagging and Search

- ▶ Very common request for OMERO to have more levels than Project and Dataset
- ▶ Driven by very reasonable need to organise data
- ▶ Why and how are tags a solution to this?
- ▶ Can tags be an improvement on this?



```
File Edit View Bookmarks Settings Help
dpwrussell@dpwrk:~/tree-structures$ tree -C date-first/Data/
date-first/Data/
├── 2010
│   ├── Collaborations
│   │   ├── John
│   │   ├── Lisa
│   │   └── Pete
│   └── 2011
│       ├── Collaborations
│       │   ├── John
│       │   └── Pete
│       └── 2012
│           ├── Collaborations
│           │   ├── John
│           │   └── Pete
│           └── 2013
│               ├── Collaborations
│               │   ├── Lisa
│               │   ├── Pete
│               │   └── DeltaVision
│               │       ├── Denoised
│               │       │   ├── File1_Info1_Info2_Info3_Denoised.dv
│               │       │   └── File1_Info1_Info2_Info3.dv
│               │   └── February
│               │       ├── DeltaVision
│               │       │   ├── Denoised
│               │       │   │   ├── File1234_Info1_Info_2_Info10_Denoised.dv
│               │       │   │   └── File1234_Info1_Info_2_Info10.dv
│               │       └── January
│               │           └── DeltaVision
│               └── 24 directories, 4 files
└── dpwrussell@dpwrk:~/tree-structures$
```

Figure: Filesystem Hierarchy



```
dpwrussell@dpwrwrk:~/tree-structures$ tree -C collab-first/Data/
collab-first/Data/
├── 2010
├── 2011
├── 2012
├── 2013
├── February
│   ├── DeltaVision
│   │   └── Denoised
│   │       ├── File1234_Info1_Info2_Info10_Denoised.dv
│   │       └── File1234_Info1_Info2_Info10.dv
│   └── January
├── Collaborations
│   ├── John
│   │   ├── 2010
│   │   ├── 2011
│   │   └── 2012
│   ├── Lisa
│   │   ├── 2010
│   │   └── 2013
│   └── Pete
│       ├── 2010
│       ├── 2011
│       ├── 2012
│       ├── 2013
│       └── DeltaVision
│           └── Denoised
│               ├── File1_Info1_Info2_Info3_Denoised.dv
│               └── File1_Info1_Info2_Info3.dv
└── 23 directories, 4 files
dpwrussell@dpwrwrk:~/tree-structures$
dpwrussell@dpwrwrk:~/tree-structures$
```

Figure: Alternate Filesystem Hierarchy

The Important Information

- ▶ E.g. 'File1_Info1_Info2_Info3_Denoised.dv'
- ▶ From directory structure: **2013, Collaborations, Pete, DeltaVision, Denoised**
- ▶ From filename: **Info1, Info2, Info3**

Using Tags Instead

- ▶ Assign tags for all important information
- ▶ The only difference in navigating tagged data as opposed to a filesystem is greater choice of path to take



Figure: All Tags, Searchable



Figure: Selecting Starting Point



Figure: Initial Tag Intersection and Results

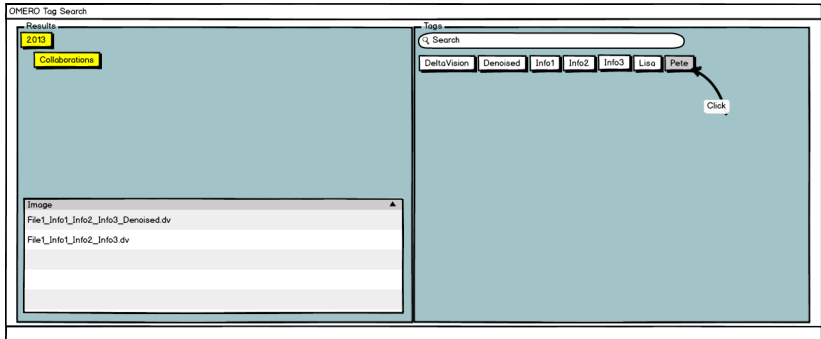


Figure: Reduced Tag Intersection and Results



Figure: Further Reduced Tag Intersection and Results



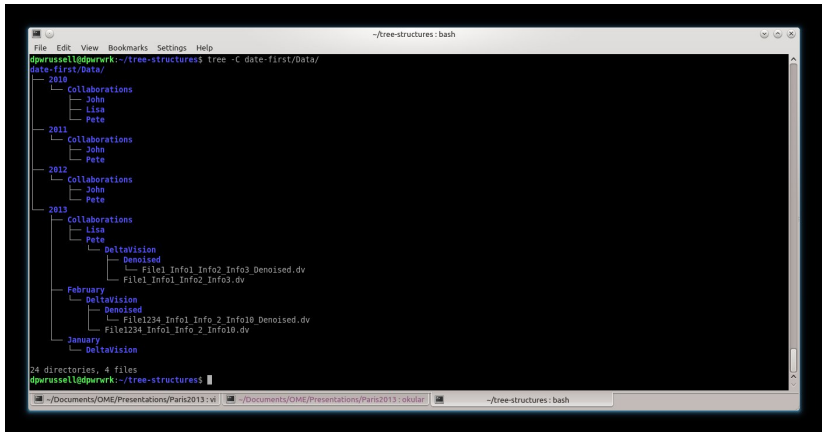
Figure: Further Reduced Tag Intersection and Results



Figure: Exact Replica of Filesystem Hierarchy

Nested Common Search Criteria

- ▶ What if the criteria of interest is not at the top level of the filesystem tree?
- ▶ Difficult to search in a filesystem.
- ▶ What about with tags?
- ▶ E.g. All Denoised data from 2013



```
File Edit View Bookmarks Settings Help
dpwrussell@dpwrk:~/tree-structures$ tree -C date-first/Data/
date-first/Data/
├── 2010
│   ├── Collaborations
│   │   ├── John
│   │   ├── Lisa
│   │   └── Pete
│   └── 2011
│       ├── Collaborations
│       │   ├── John
│       │   └── Pete
│       └── 2012
│           ├── Collaborations
│           │   ├── John
│           │   └── Pete
│           └── 2013
│               ├── Collaborations
│               │   ├── Lisa
│               │   ├── Pete
│               │   └── DeltaVision
│               │       ├── Denoised
│               │       │   ├── File1_Info1_Info2_Info3_Denoised.dv
│               │       │   └── File1_Info1_Info2_Info3.dv
│               └── February
│                   ├── DeltaVision
│                   │   ├── Denoised
│                   │   │   ├── File1234_Info1_Info_2_Info10_Denoised.dv
│                   │   │   └── File1234_Info1_Info_2_Info10.dv
│                   └── January
│                       └── DeltaVision
24 directories, 4 files
dpwrussell@dpwrk:~/tree-structures$
```

Figure: Filesystem Hierarchy

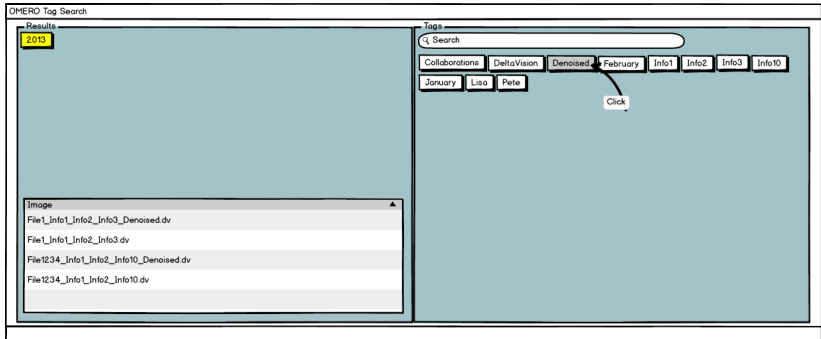


Figure: '2013' Starting Point



Figure: Results matching '2013' and 'Denoised'

In Practice

- ▶ Mockups based on design discussion with Will, Jean-Marie, Gus and others
- ▶ For implementation in OMERO Clients
- ▶ Search needs some work