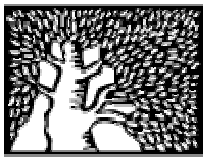
BioImg Storage Server

Ofra Golani, Kirill Kogan, Jaime Prilusky



מכון ויצמן למדע
WEIZMANN INSTITUTE OF SCIENCE

BioImg Overview

Central repository for raw imaging data that provide service for all core facilities at the Weizmann Institute

Data is copied automatically without conversion from all imaging resources

Users can **document, search, access** and **share** data from anywhere on campus

The BioImaging Community at WIS

Distributed and Diverse facilities

- Diverse imaging techniques
- Distributed imaging facilities
- Many Instrument's administrators
- Diverse image visualization and analysis tools

Common Problems

- Hard to find data
- Data storage on the instruments got full quickly
- Use of Disk-on-keys $\Rightarrow$ Computer Viruses
- Data duplication on many computers

BioImg Objectives

To Improve **storage and retrieval of data** from existing imaging infrastructure at the core service, departmental facilities and individual research lab level.

To set up an **annotation interface**, to add metadata, enabling **organized and effective searching and retrieving**, and improved management of infrastructure and data.

BioImg is not a backup system.

Main Features

- Central repository for raw Imaging data.
- Data is copied automatically from all connected instruments and users – No user initiation is needed
- Data is saved as is - No file conversion is applied
- Data is organized on *BioImg* by Instruments, users, and experiments
- Users can document, search, access and share data from anywhere on campus.
- Use your favorite visualization and analysis tools - No built-in Analysis tools.
- Analyzed data is not saved on *BioImg*
- Authentication is done by the institute's LDAP system
- Data is automatically shared with the PI
- Data can be shared with people out of campus using http-based service

BioImg Statistics

- ~60 instruments
  - Core facilities, Departmental Microscopes, Lab microscopes
  - MRI, CT, Electron Microscopy, light microscopy, Digital Histology, FACS
- 1000+ users
- Data: 233TB (out of 544TB) + Mirror
- Over 5 years

What *Bioimg* offers to users ?

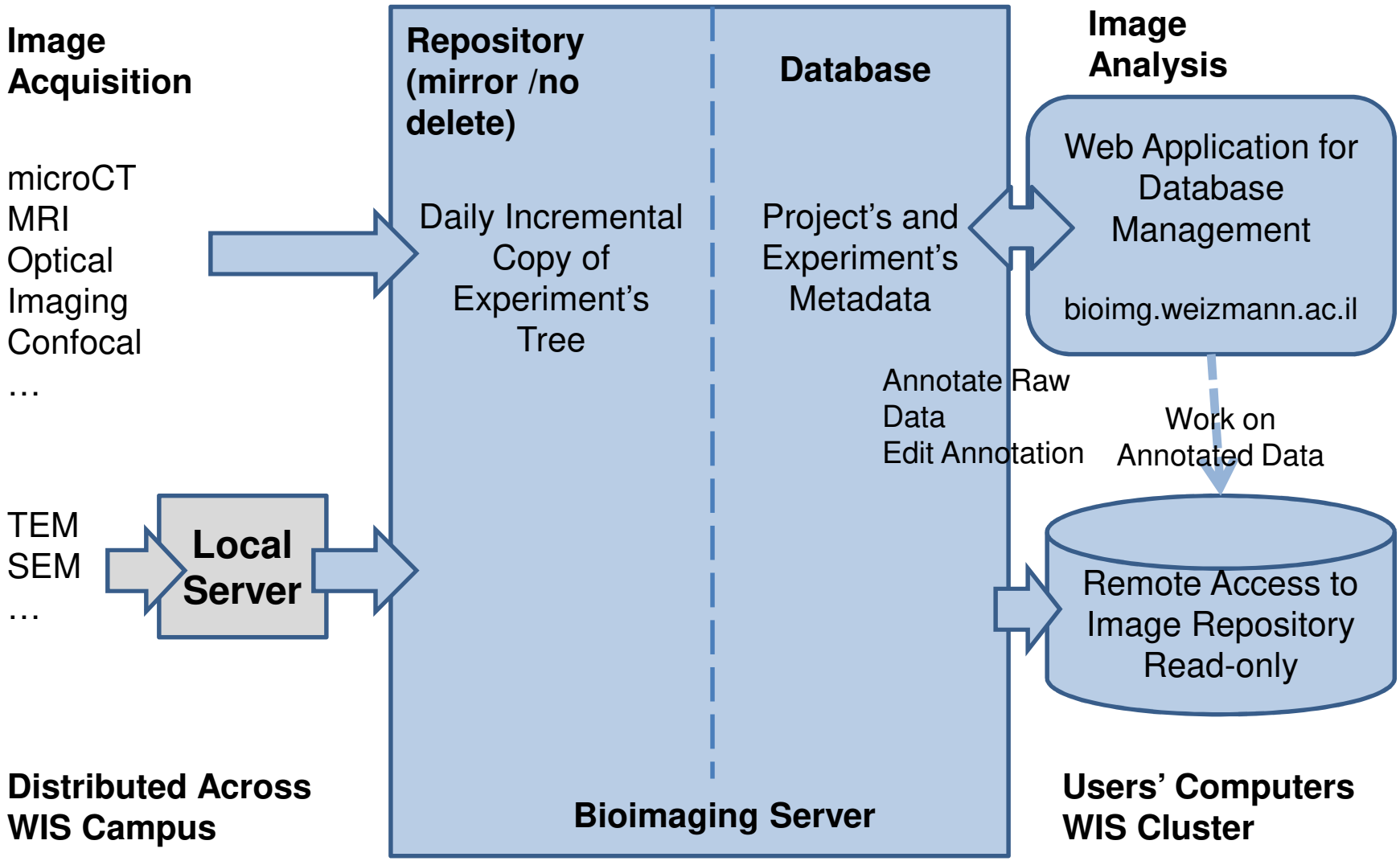
Simplicity: single interface and access for all equipment
Access: anywhere over network, inside and outside WIS
Privacy: protected access to own data
Sharing: shared access to experiments and projects
Export: send experiments over email
Documentation: save notes with the data
Search: locate data by user, notes and date
Personalization: own preferences tailor Bioimg behavior
Awareness: email notification on availability of new data
Actualization: data copied up to 8 times a day
Organization: PIs have access to all their group data

But …

Analyzed data not included

Data is not archived, BioImg is not a backup system

BioImg Structure



Data Organization and Access

- Data is stored by Instruments, Users and experiment
- Each experiment is a folder that can contain any type of files or sub-folders
- Main entity is experiment, experiments are logically organized by Projects
- Logical links to data are created from user folder
- Data Access by mounting the private user folder
- User's view:
  - Choose to organize by Project or Instrument
  - Choose which experiments to show, others are hidden
  - Search experiments by: Date, Project, Textual documentation
- Access, Search and Annotations at the experiment level
- User's folders are read only

Quick Start for Users

At the microscope

- How to store data at the microscope ?

Data Annotation through the website

- Assign the experiment to a project (must)
- Add Notes (optional, highly recommended)

Data Access

- Mount your virtual disk on your personal computer
- Share with other users in WIS
- Send by mail to other users

Storing Data on the Microscope

- Keep the data in the BioImg folder under your folder

  D:\DATA_ON_BIOIMG\userid
- Data should be saved in sub-folders.
- Experiment = folder under the userid folder
- Don’t save single data files in the userid directories
- Keep the folder structure as flat as possible.
- Rule of thumb: **create a new folder whenever you start a new page in you lab notebook**.
- Use meaningful names: Cond_Date

  D:\DATA_ON_BIOIMG\ofrag\20140105_Demo_Exp

Storing Data on the Microscope

- Allowed characters in folders/files names are:

  A-Z a-z 0-9 _-.
- Don’t use special characters in folder names, ie:

  ‘,”`!@#$%^&?()[]{}<>*+=/\
- Don’t change folder/file name
- Store only raw files and not processed data.
- You can put short text file with description (readme.txt)

What about old data ?

- Organize the Old data in the new format: folder per experiment
- Use “Move” and not “Copy” to preserve the original dates and to save space

Quick Start for Users

At the microscope

- How to store data at the microscope ?

Data Annotation through the website

- Assign the experiment to a project (must)
- Add Notes (optional, highly recommended)

Data Access

- Mount your virtual disk on your personal computer
- Share with other users in WIS
- Send by mail to other users

Experiments and Projects

- ***Experiment* is a directory** of image files, taken at a certain time, with a certain instrument, **located at the highest level in the userid folder**.
- Each ***Experiment*** is assigned to a ***Project***.
- A project can have many experiments, from one or more instruments.
- Add notes at the Experiment and the Project level (*Annotation*).
- Search data by Experiment and Project name/notes/date
- Access data at the experiment level.
- An Experiment can have sub directories but annotation is NOT possible at the sub directories level.

Quick Start for Users

At the microscope

- How to store data at the microscope?
- Ask microscope Administrator.

Data Annotation through the website

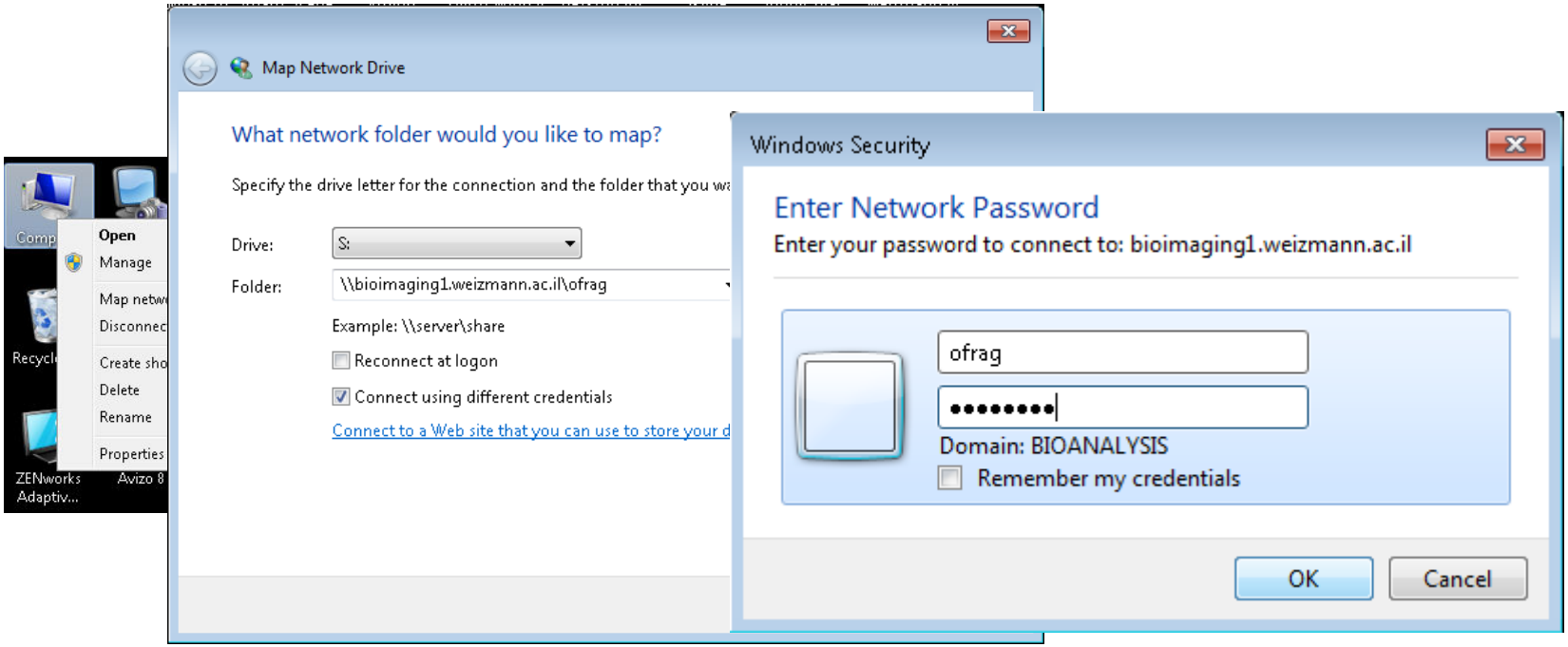
- Assign the experiment to a project (must)
- Add Notes (optional, highly recommended)

Data Access

- Mount your virtual disk on your personal computer
- Share with other users in WIS
- Send by mail to other users

Connect your Virtual Disk

PCs: use *Map Network Drive*

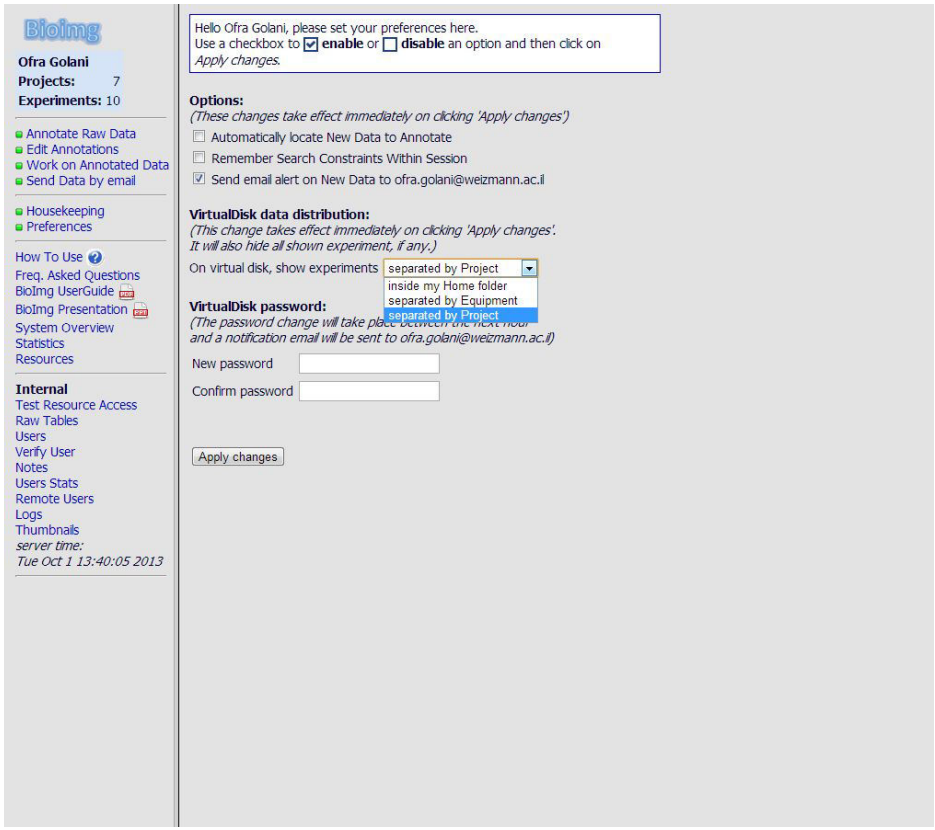


Default password is ‘YourUserid1’
Change it using the *Preference* page on the website

Mac: use *Connect to Server*

Control your Virtual Disk

- Show / Hide data using ‘Work on annotated data’
- Control organization using ‘Preferences



Troubleshooting

- I can’t connect to my virtual disk
  - No WiFi
  - Default password is ‘YourUserid1’
- My data was not copied
  - Verify that you did not use special characters
- I got an email notification, but I can’t annotate my data
  - If you put new data in existing subfolder you only get email but don’t need to annotate

Summary

- At the microscope:
  - Create new folder in D:\UserData\youruerid for each experiment
- bioimg.weizmann.ac.il
  - *"*Annotate Raw Data*"* to assign the experiment to a project (must)
  - "Edit Annotation" to add notes (recommended)
- Data Access
  - Mount your virtual disk on your personal computer
  - Share with other users (in/out of WIS)
- Make a second copy for archive if needed

Getting Help

- Bioimg.weizmann.ac.il
  - How To Use
  - Frequently Asked Questions
  - User Guide
  - This Presentation
- Instruction sheet at the microscope
- Ofra.Golani@weizmann.ac.il / x5177
- Kiril.Kogan@weizmann.ac.il / x3113
- Jaime.prilusky@weizmann.ac.il / x4959

Bioimg for Facility Administrators

Serenity: No Need to worry if users downloaded their data and about acquiring more hard drives

Security: No more disk-on-keys

Access: Administrators can access all the data acquired on the instrument

Sharing: Easy way to share data they acquire for users

Organization: Data is organized by userids in a common away across all instruments

Reliability: RAID, mirror

Bioimg for System Administrators

Adding New Instruments

Instrument side: creating folders, sharing, setting firewall exceptions

***BioImg* side:** simple setting using the browser based application

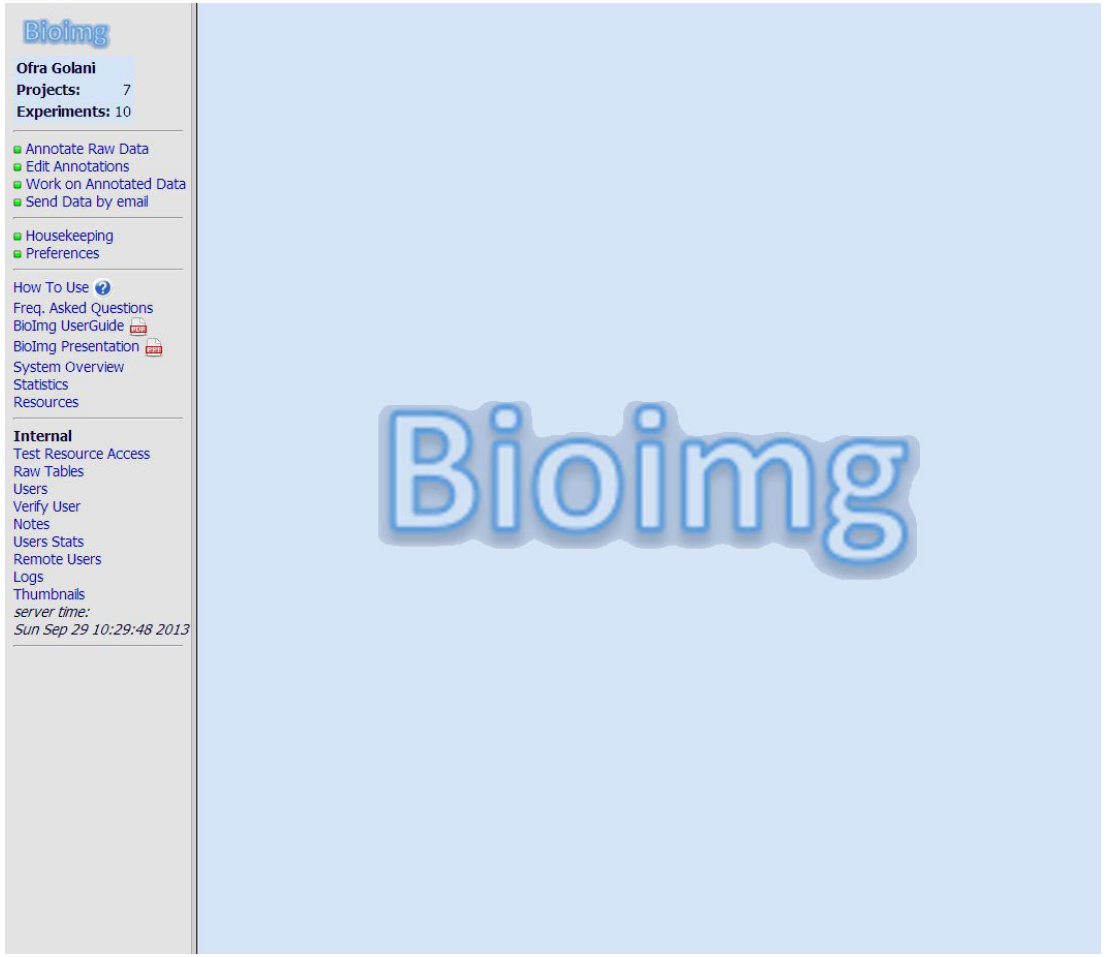
Adding New Users

Mostly automatic, an option for manual addition of users using the website

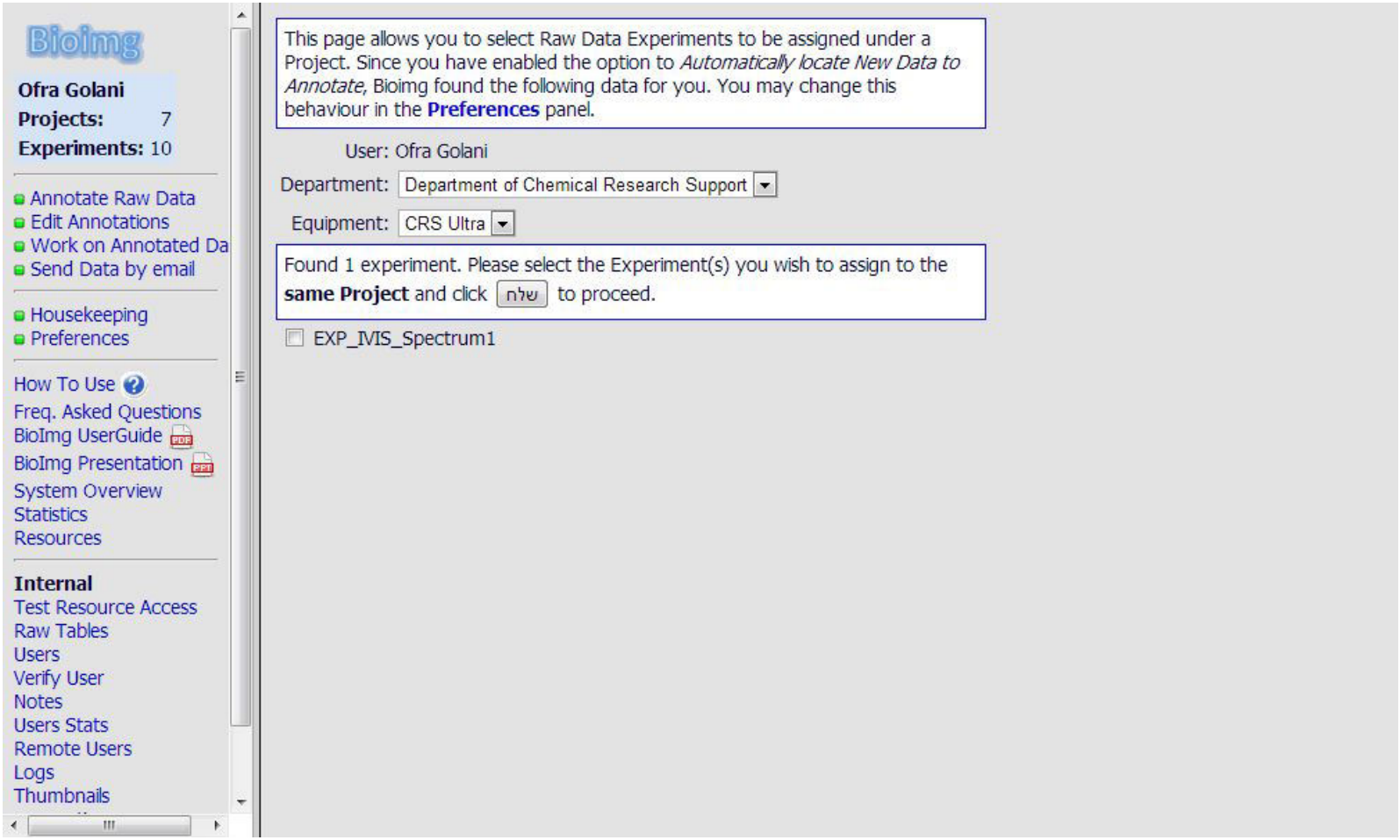
The website allows also

- Testing resource access
- Statistics
- Activity Log

User Interface

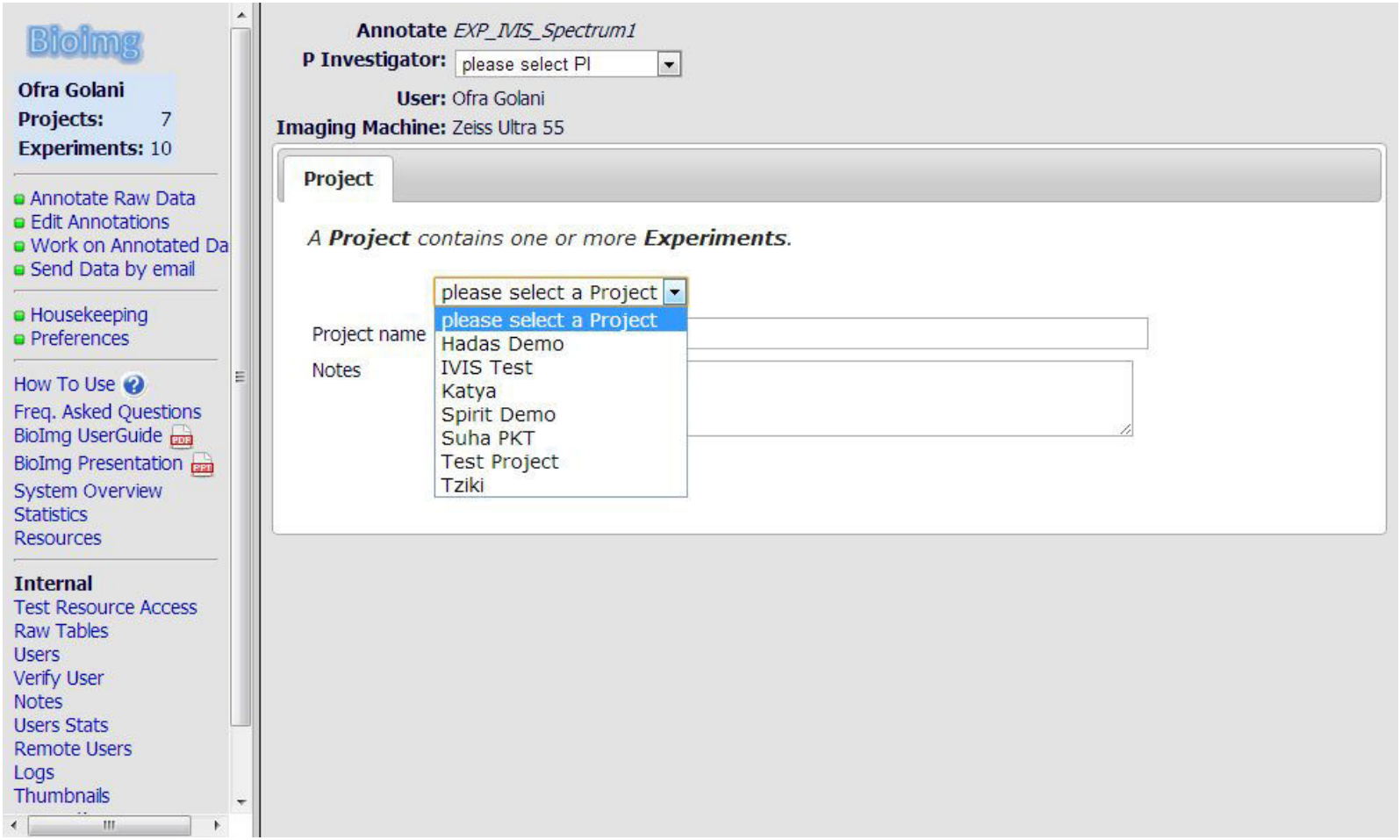


Annotate Raw Data



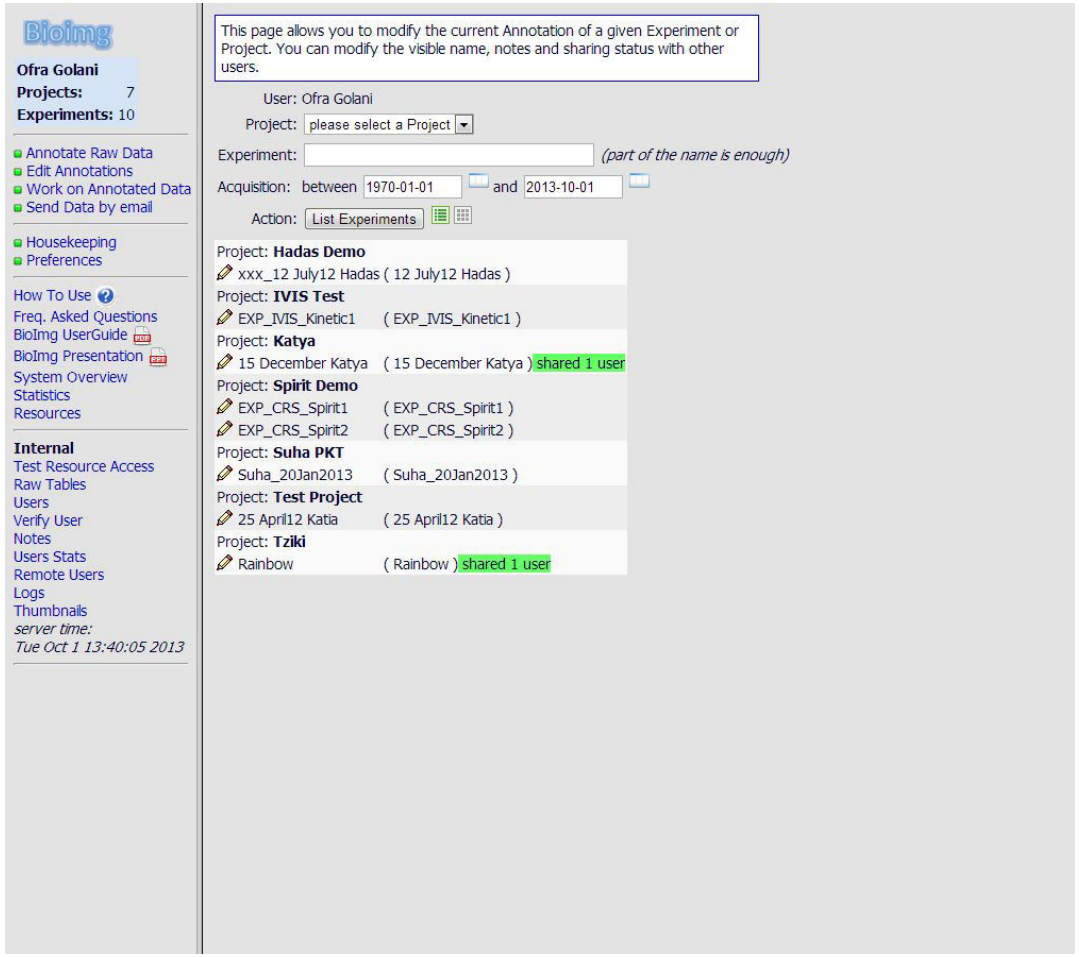
New Experiments are Listed

Annotate Raw Data



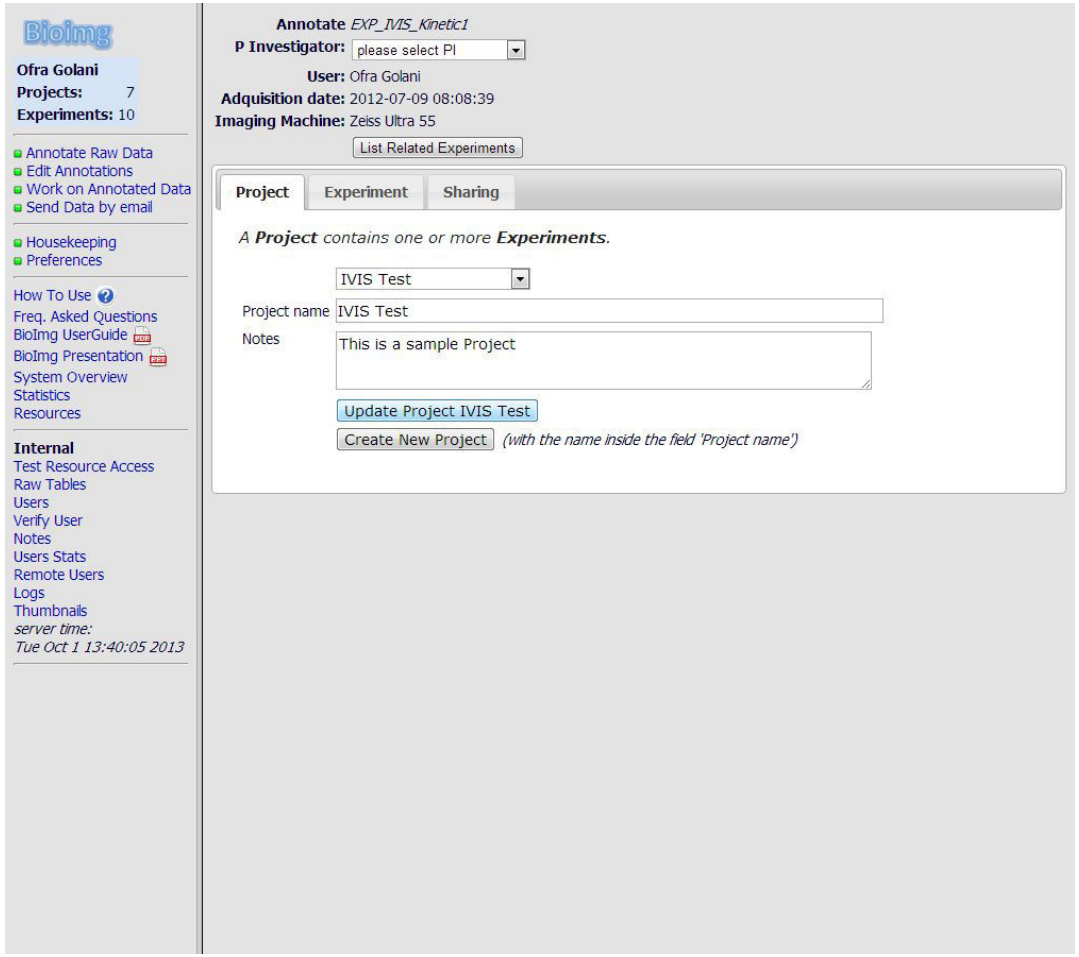
Add Experiment to the database by assigning to project

Edit Annotations



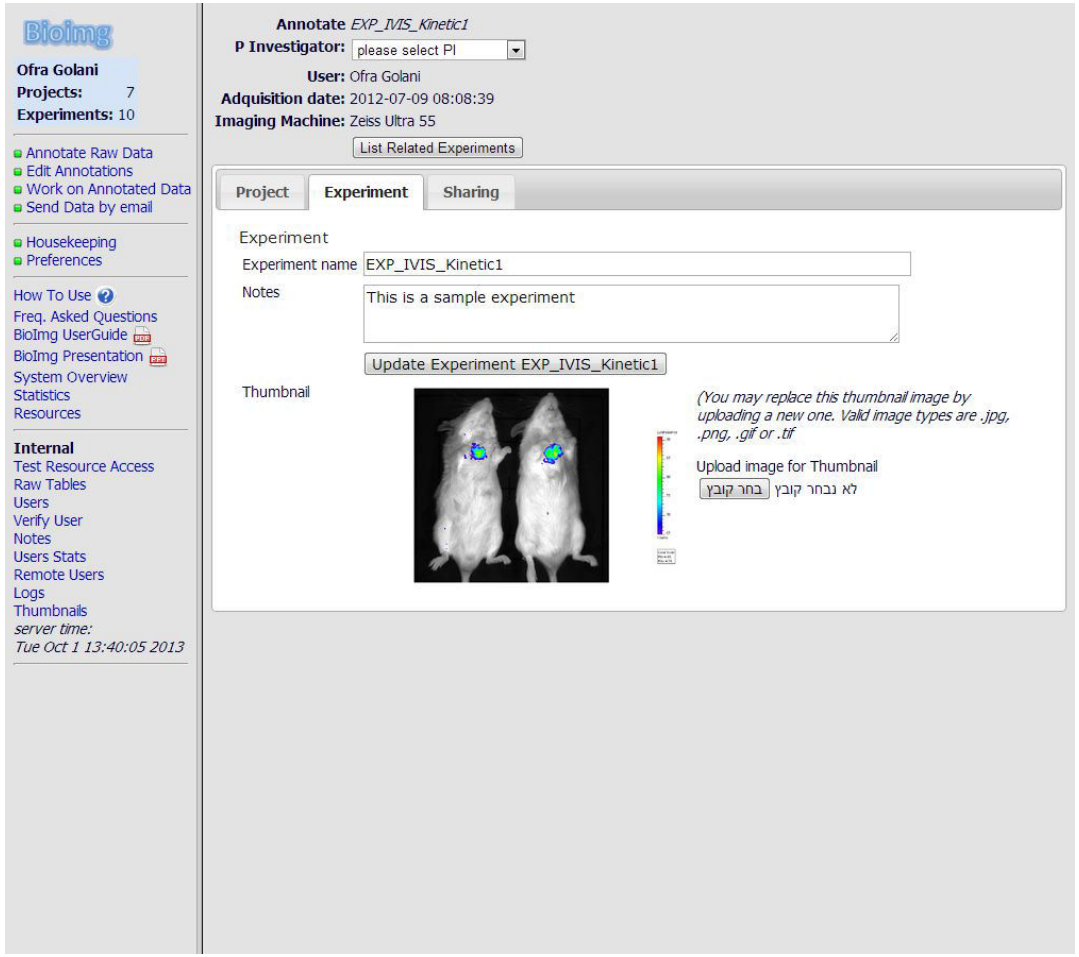
Select experiment to modify it's metadata

Edit Annotations



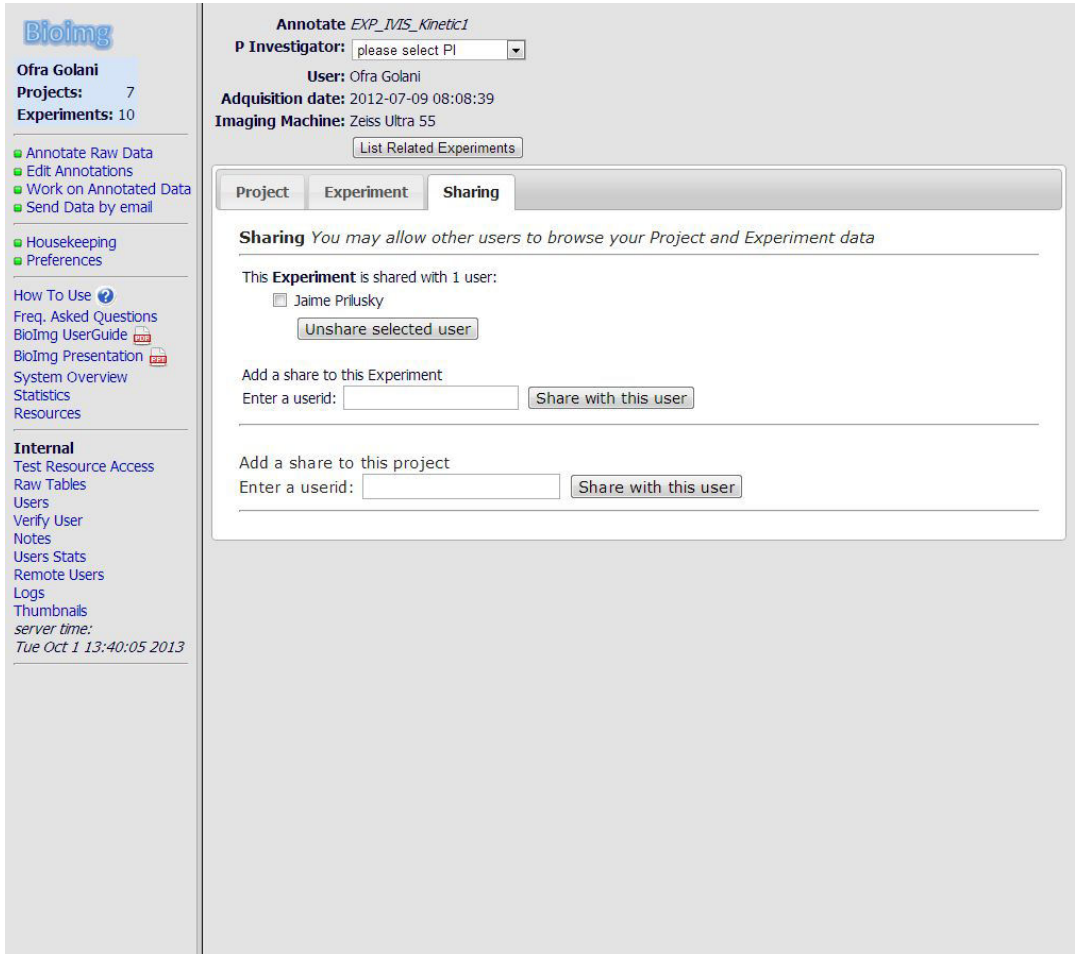
User may annotate Project additional information

Edit Annotations



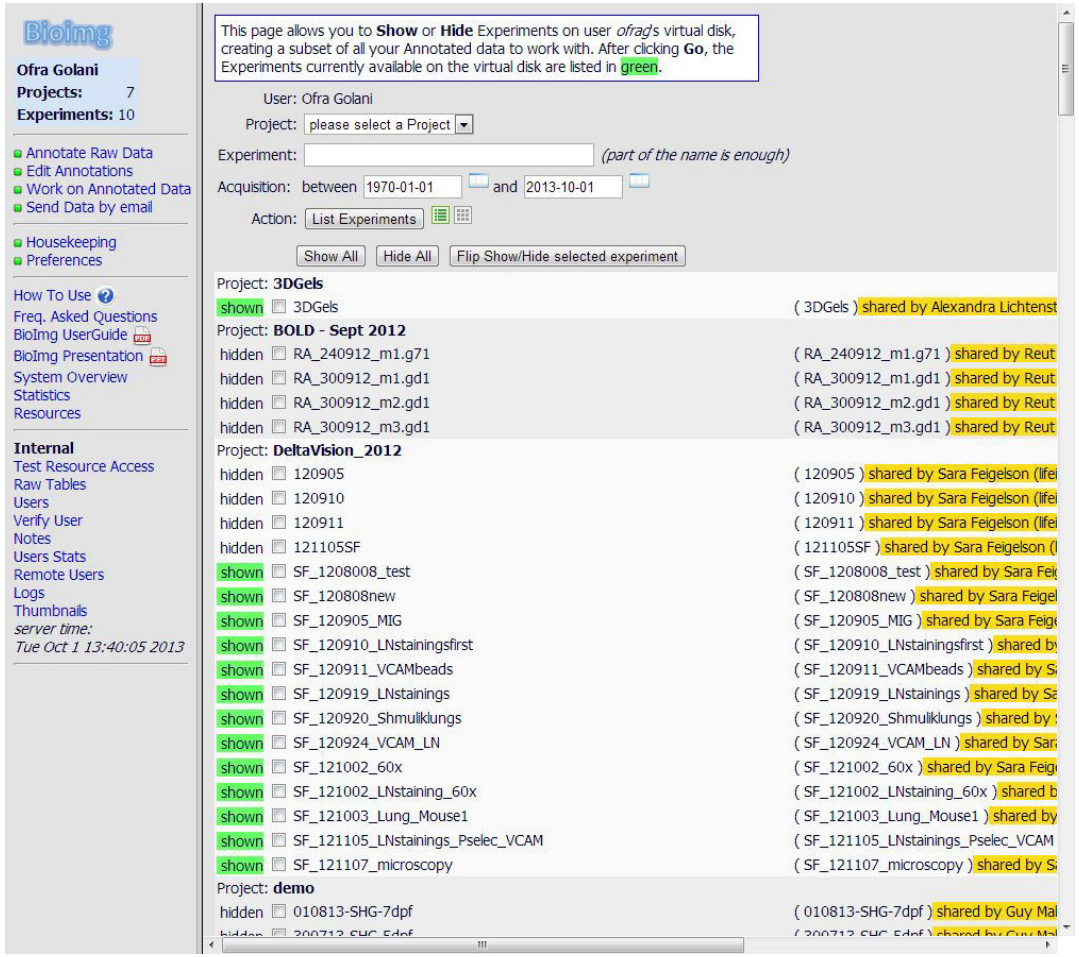
User may annotate Experiment additional information

Sharing



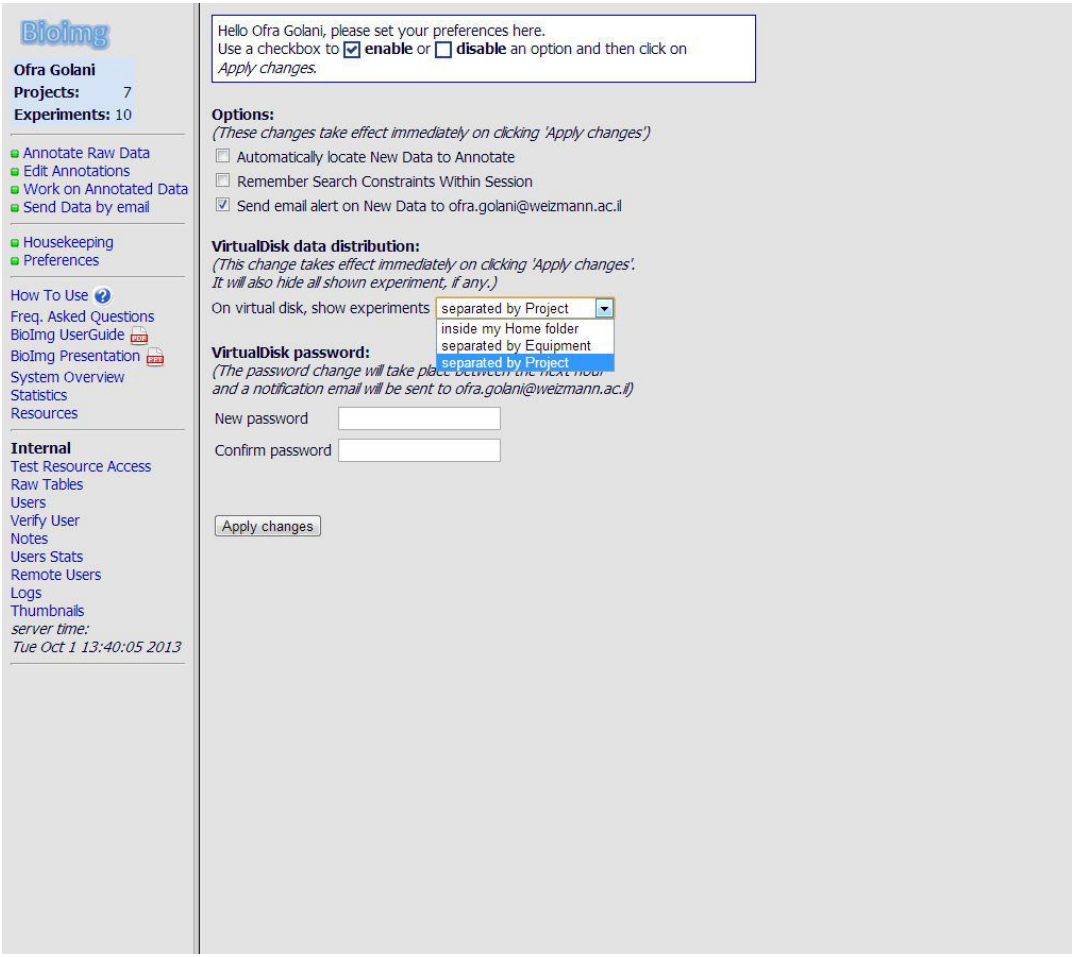
Share Project or Experiments with selected Users

Work on Annotated Data



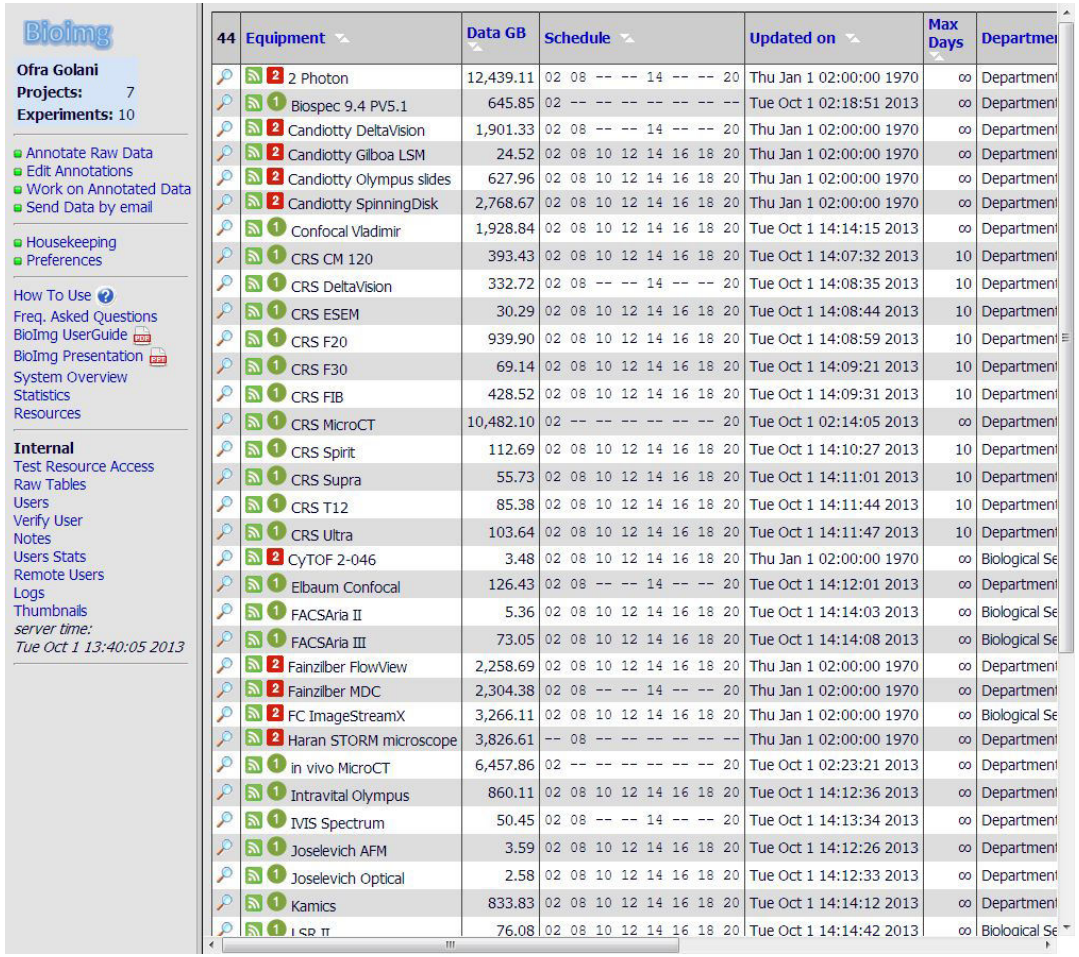
Show / Hide data on your virtual disk

Preferences



Set Preferences

Resources



Instruments information is listed and can be edited

Getting Help

- On the Website
  - How To Use
  - Frequently Asked Questions
  - User Guide
  - PDF Presentation
- Introduction Presentations for groups of users
- Instruction sheet at the microscope
- Help desk

Behind the Screen

- Scheduled Perl scripts for:
  - Copying data
  - Creating user folders
  - Sending email notification about new data
  - Updating statistics
- User Initiated Scripts:
  - Annotating data
  - Reorganizing user links (Project/Instrument view)
  - Updating User preferences
- Running on any Linux server
- MySQL Database
- Connect to the institute LDAP for user authentication
- Inside the institute firewall
- RAID 6