BISQUE and BISQUIK

Kristian Kvilekval

UCSB Bio-Imaging

- Bioimage.ucsb.edu
- 5 Year project UCSB, CMU, …
  - Manjunath: UCSB
  - Murphy: CMU
- Center for bio-image informatics
- UCSB: Ece, CS, Neuroscience institute
  - Retinal detachment : Fischer, Marc lab UT
  - MT Dynamics: Wilson & Fienstein

BISQUE

Bio-Image Semantic Query User Environment

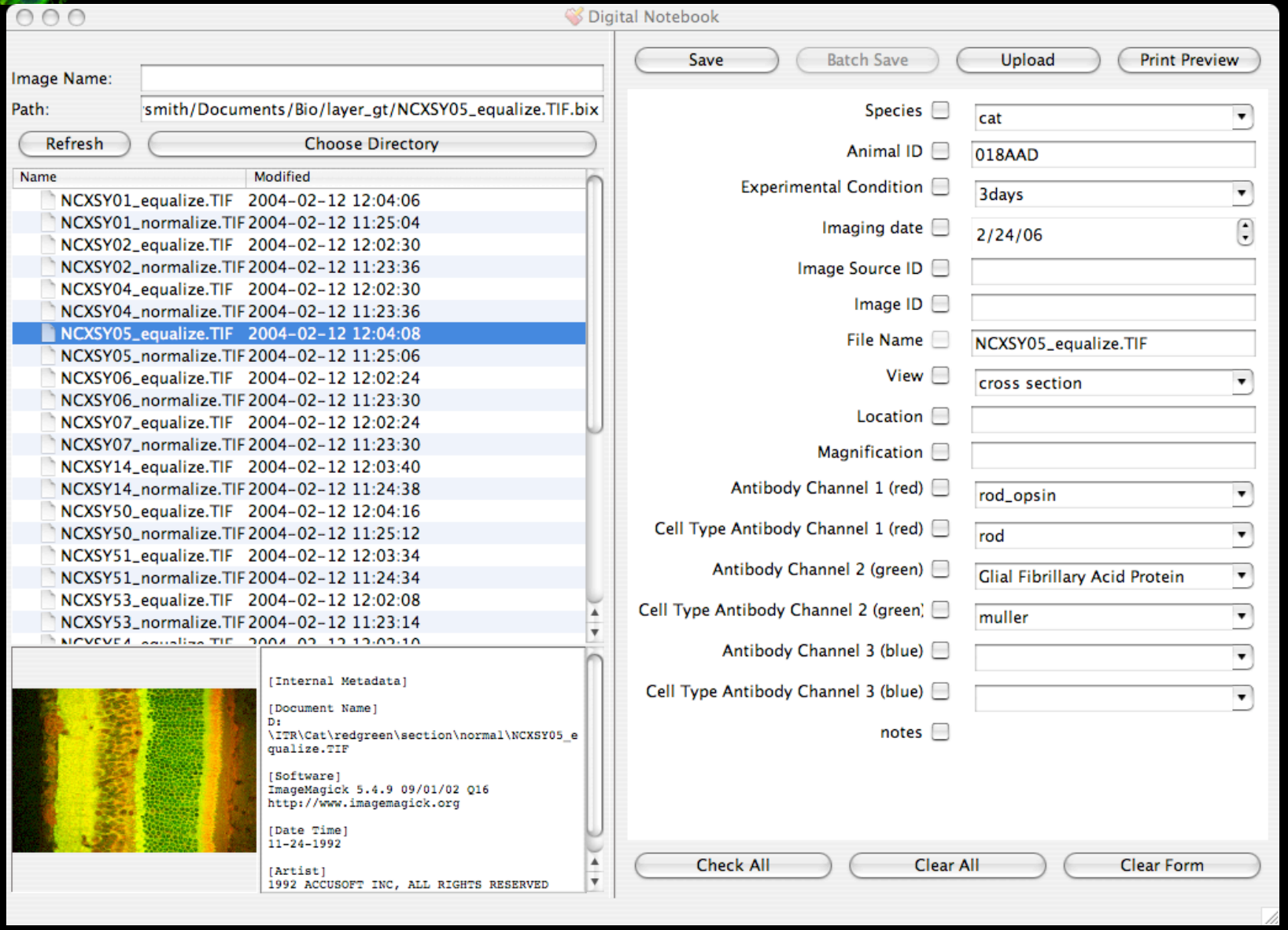
- BISQUE functionality
  - Ground truth acquisition
  - Data management
  - Analysis
- Experimental Data sets (2d images)
  - Retinal detachment experiments 50k
  - Retinal EM (cell morphology) 23k
  - Microtubule >100K

Data management capabilities

- Digital notebook
  - Capture experimental/image parameters
  - Direct import process
  - Reconfigurable (confocal, 4 x microtubule, etc)
  - Offline and Online
- Web
  - Web access and browsing
  - Organize images and metadata
  - Data sharing environment
  - Search by metadata or content
  - Integrated analysis

Center for Bioimaging Informatics

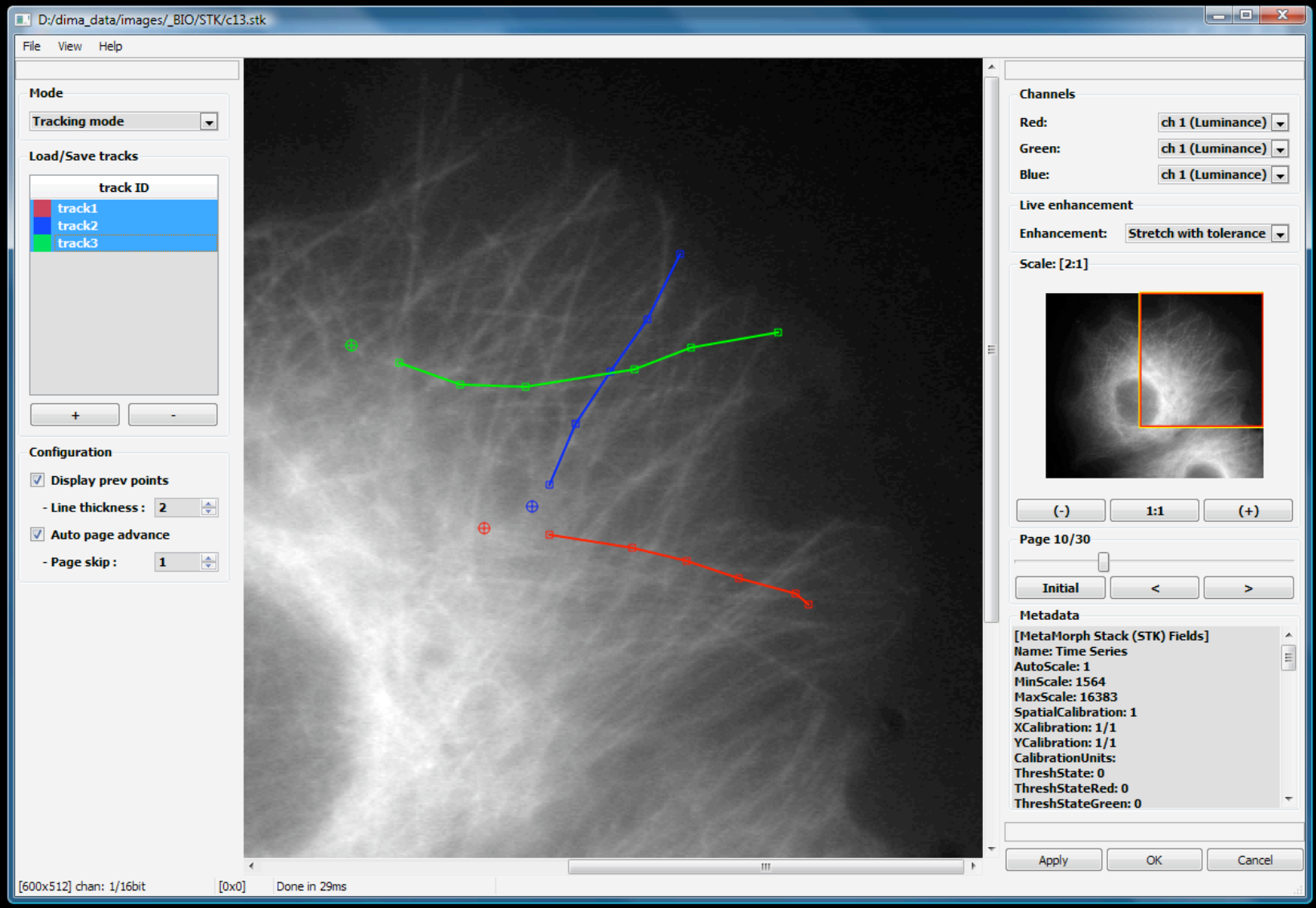
Screenshots (Digital Notebook)



May 2007

Center for Bioimaging Informatics

Graphical Ground truth



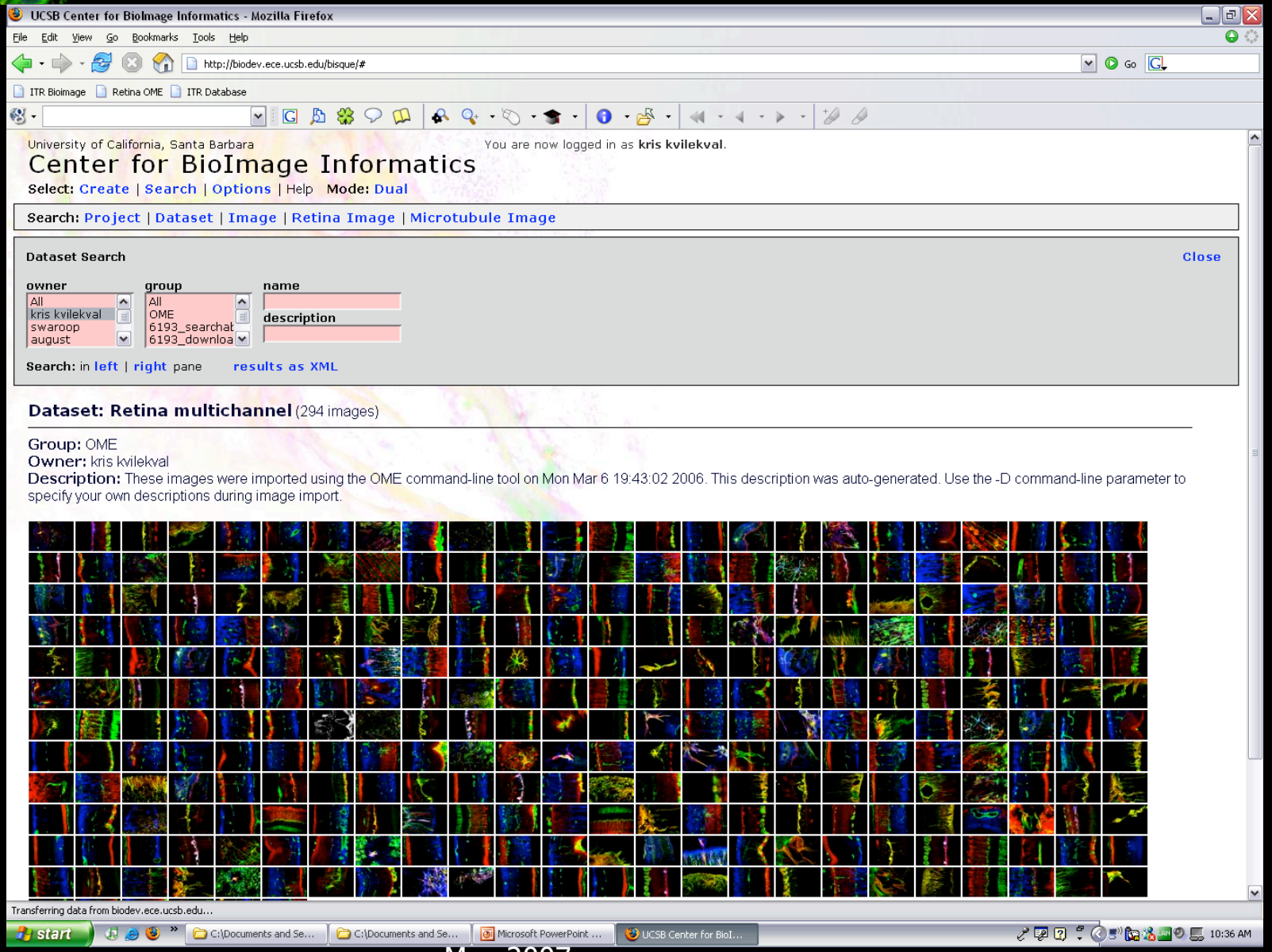
May 2007

Data management capabilities

- Web accessible features
  - Experimental image/data capture and upload
  - Image access and browsing
  - Organize images and metadata
  - Data sharing environment
  - Search by metadata or content
  - Integrated analysis

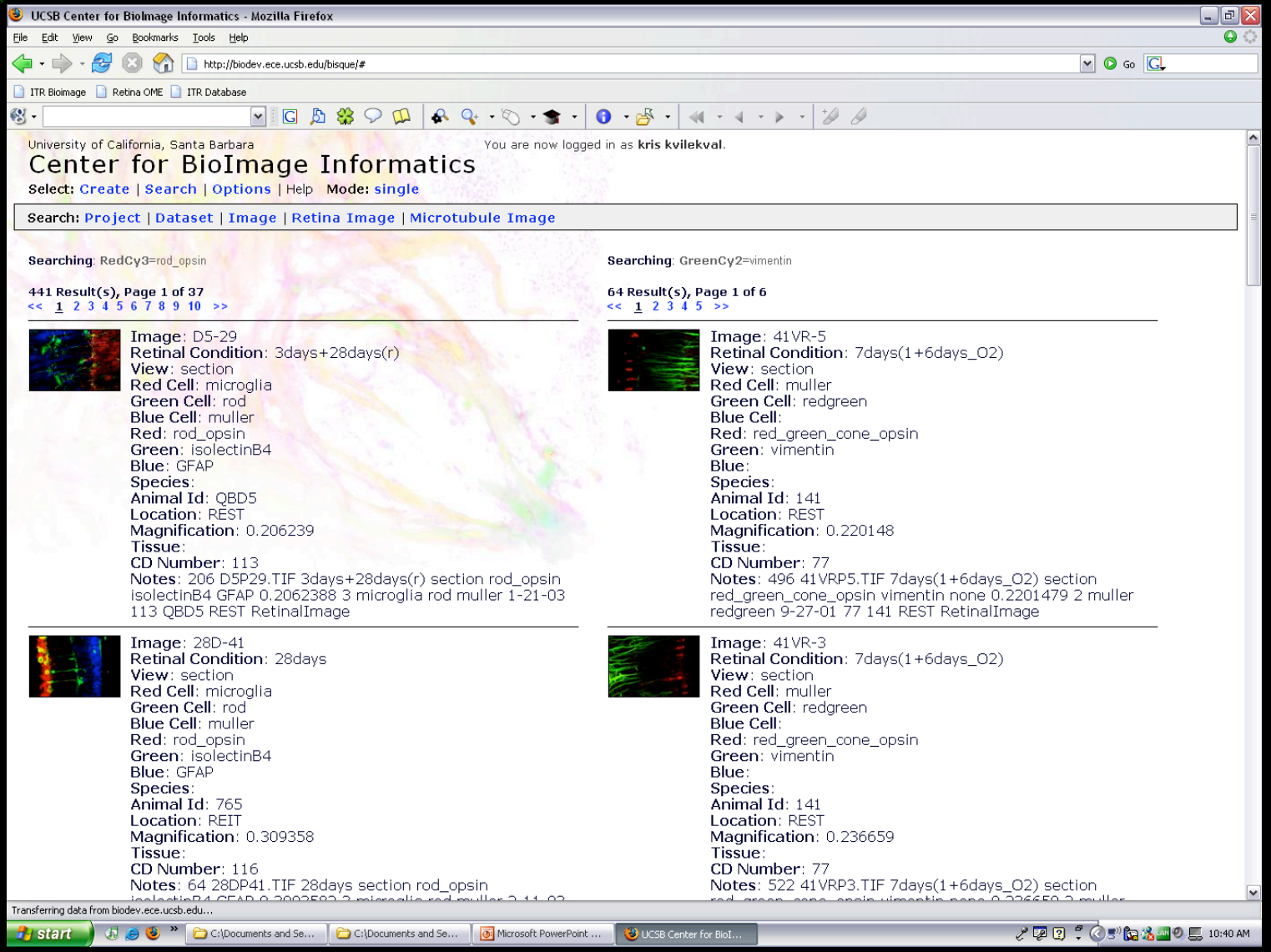
Center for Bioimaging Informatics

Screenshots (browsing)



Center for Bioimaging Informatics

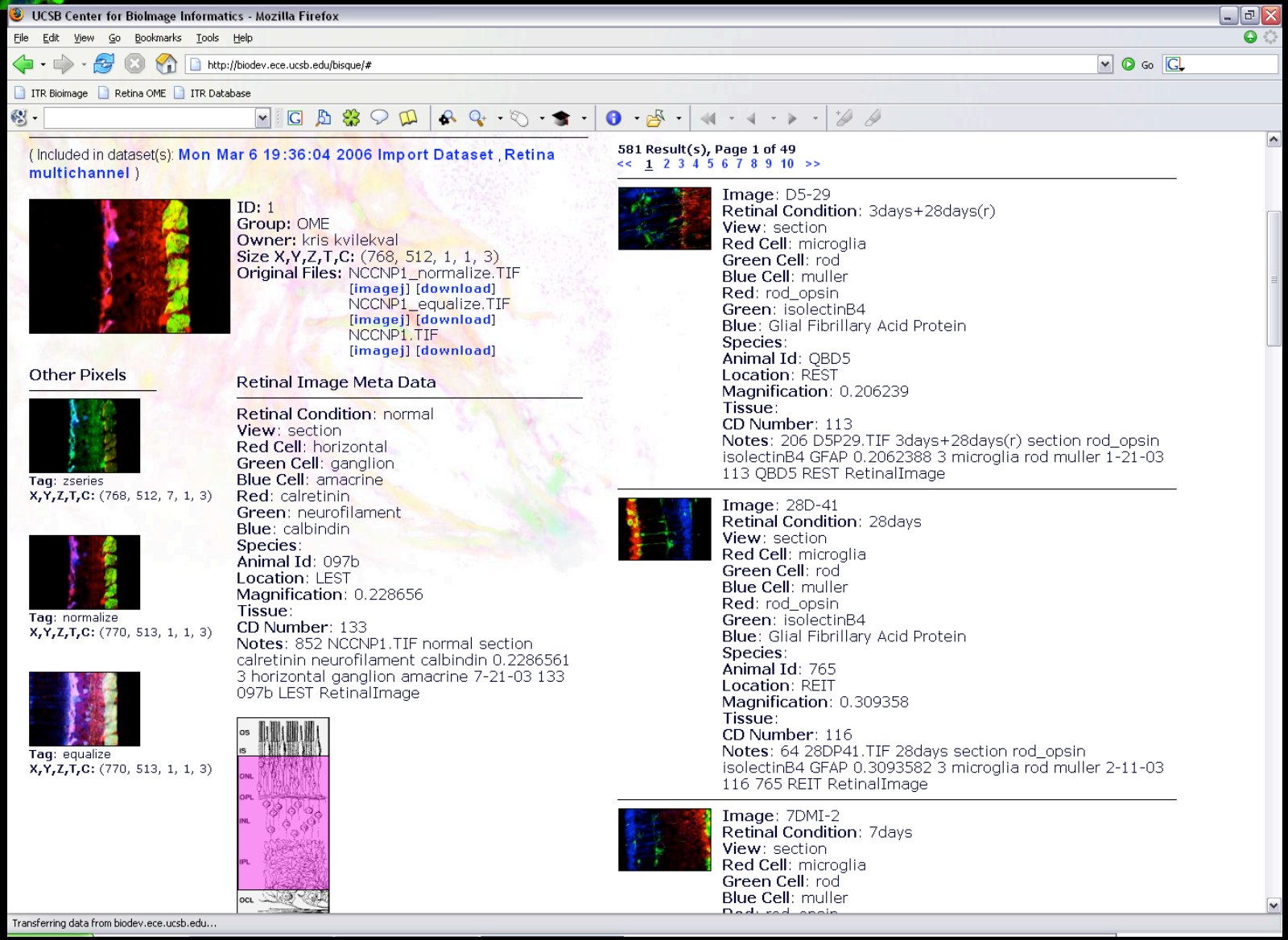
Screenshots (search)



May 2007

Center for Bioimaging Informatics

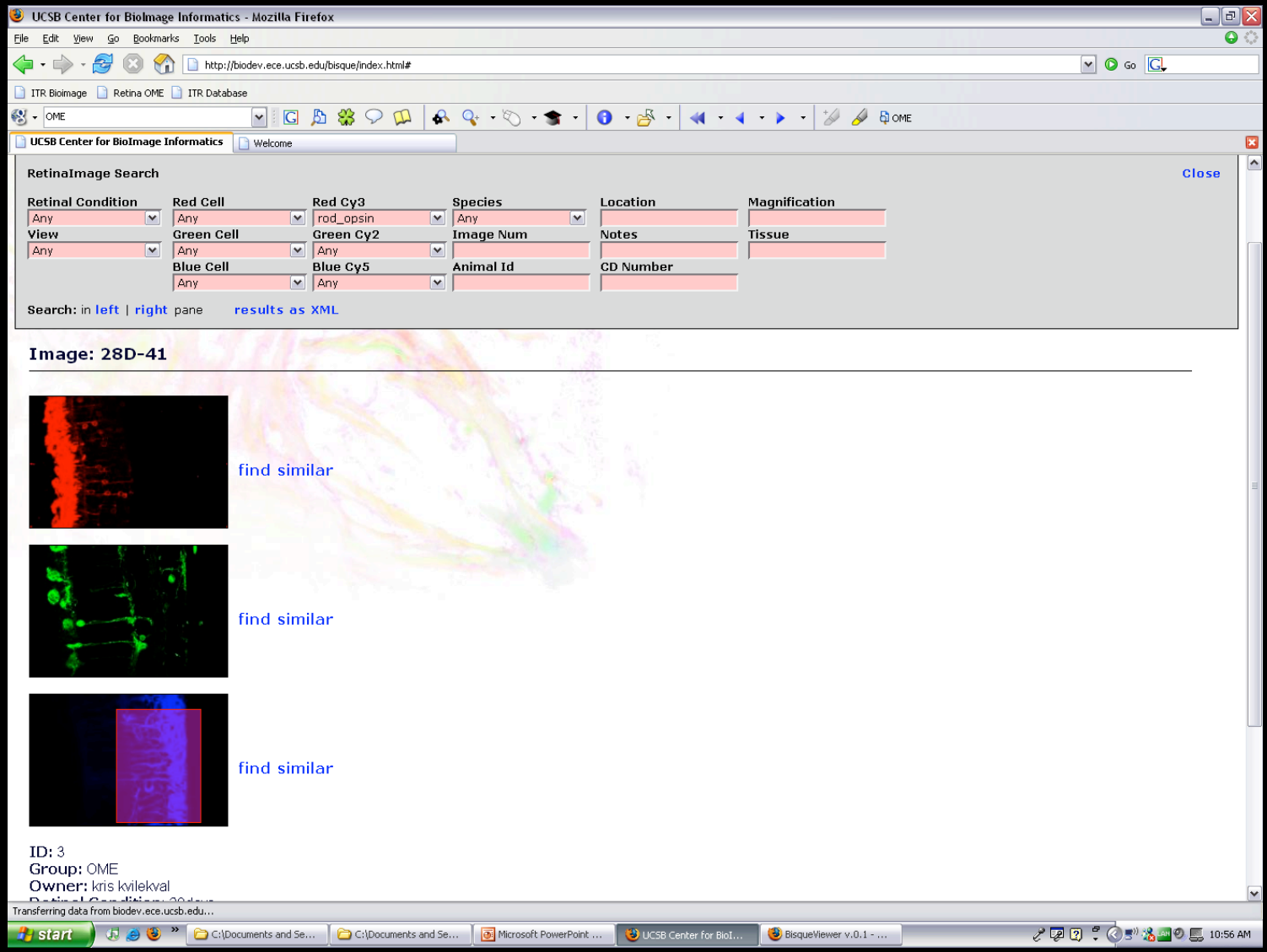
Screenshots (search)



May 2007

Center for Bioimaging Informatics

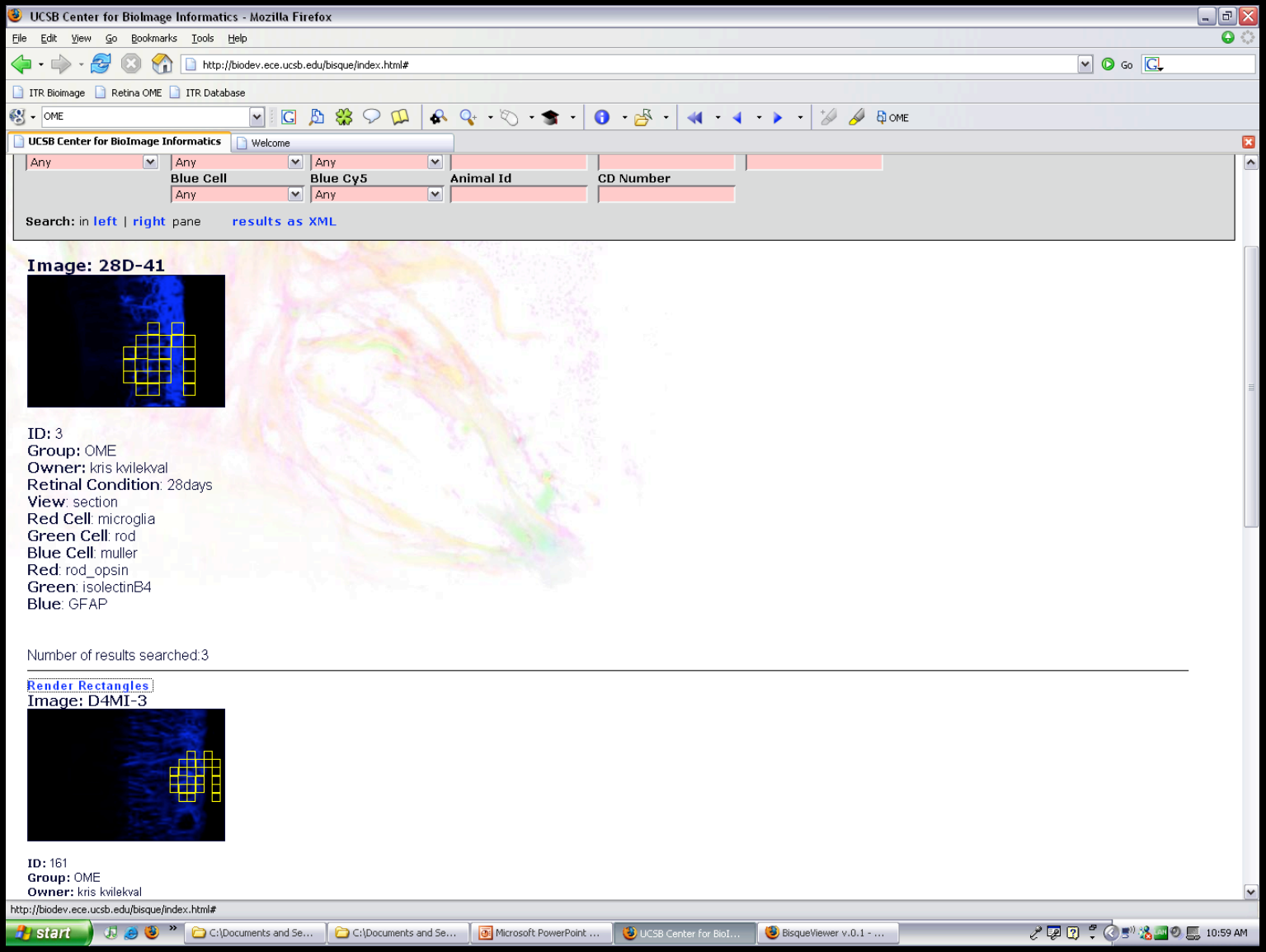
Screenshots (search similar)



May 2007

Center for Bioimaging Informatics

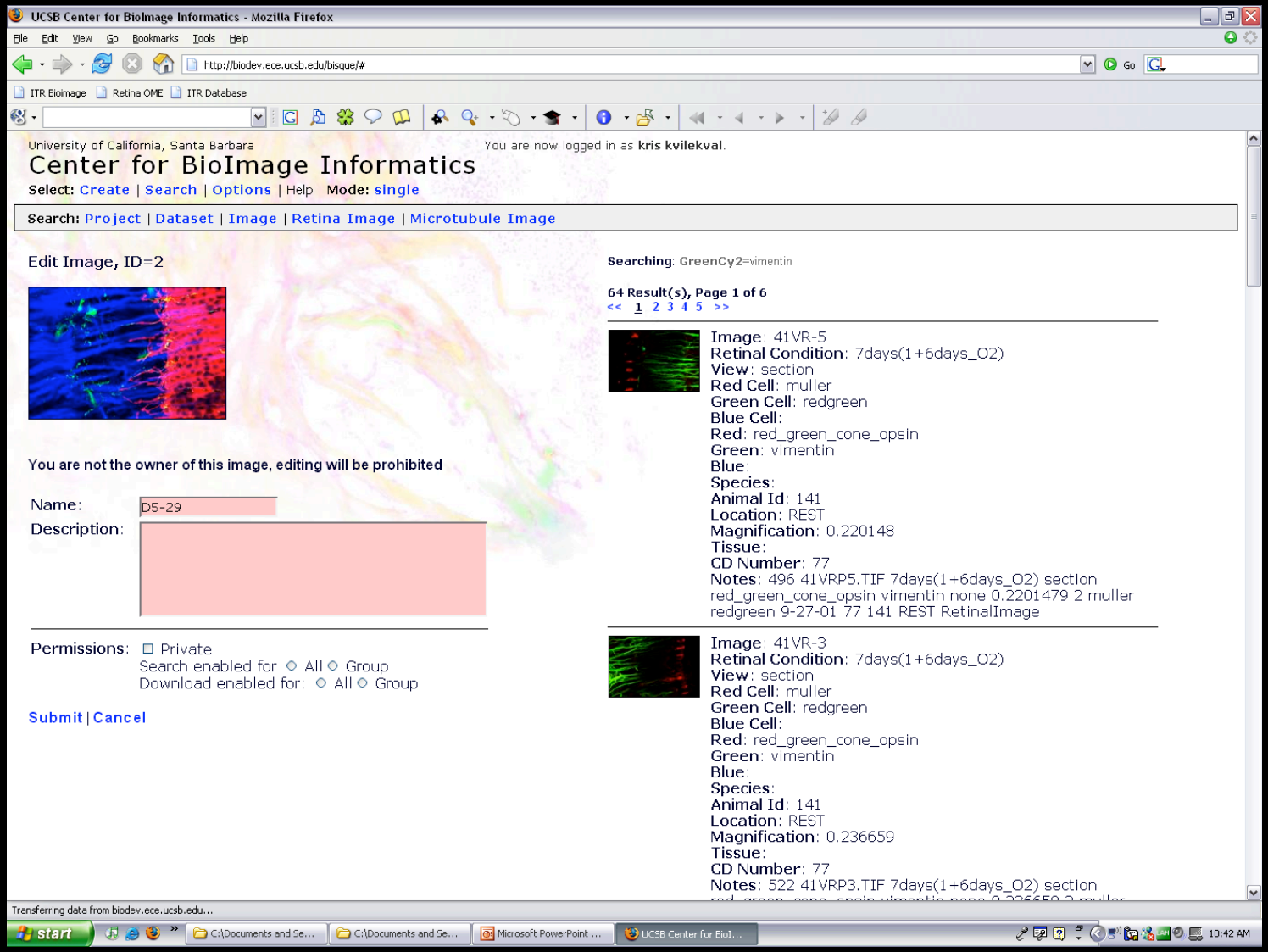
Screenshots (search similar)



May 2007

Center for Bioimaging Informatics

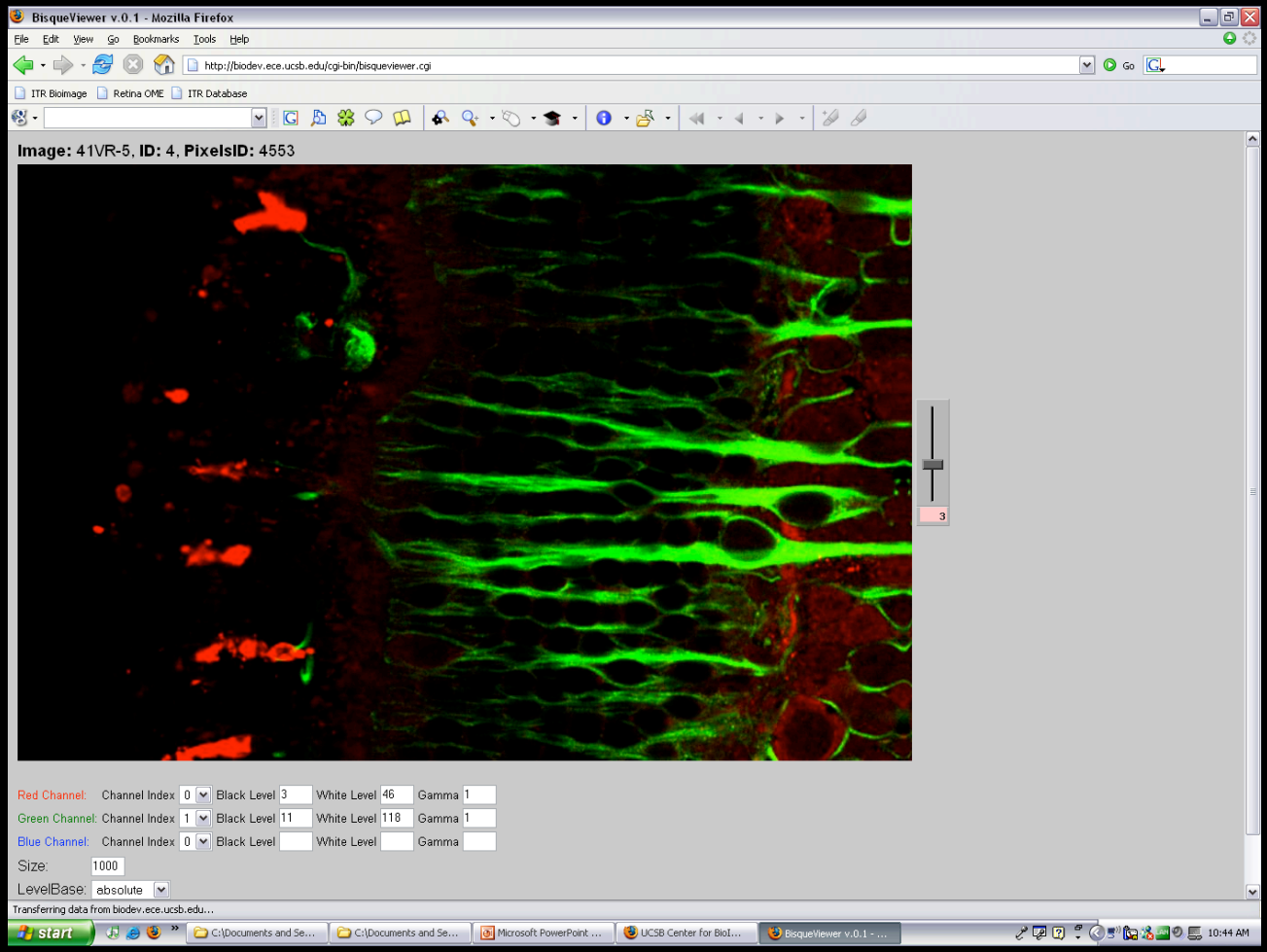
Screenshots (personal collection)



May 2007

Center for Bioimaging Informatics

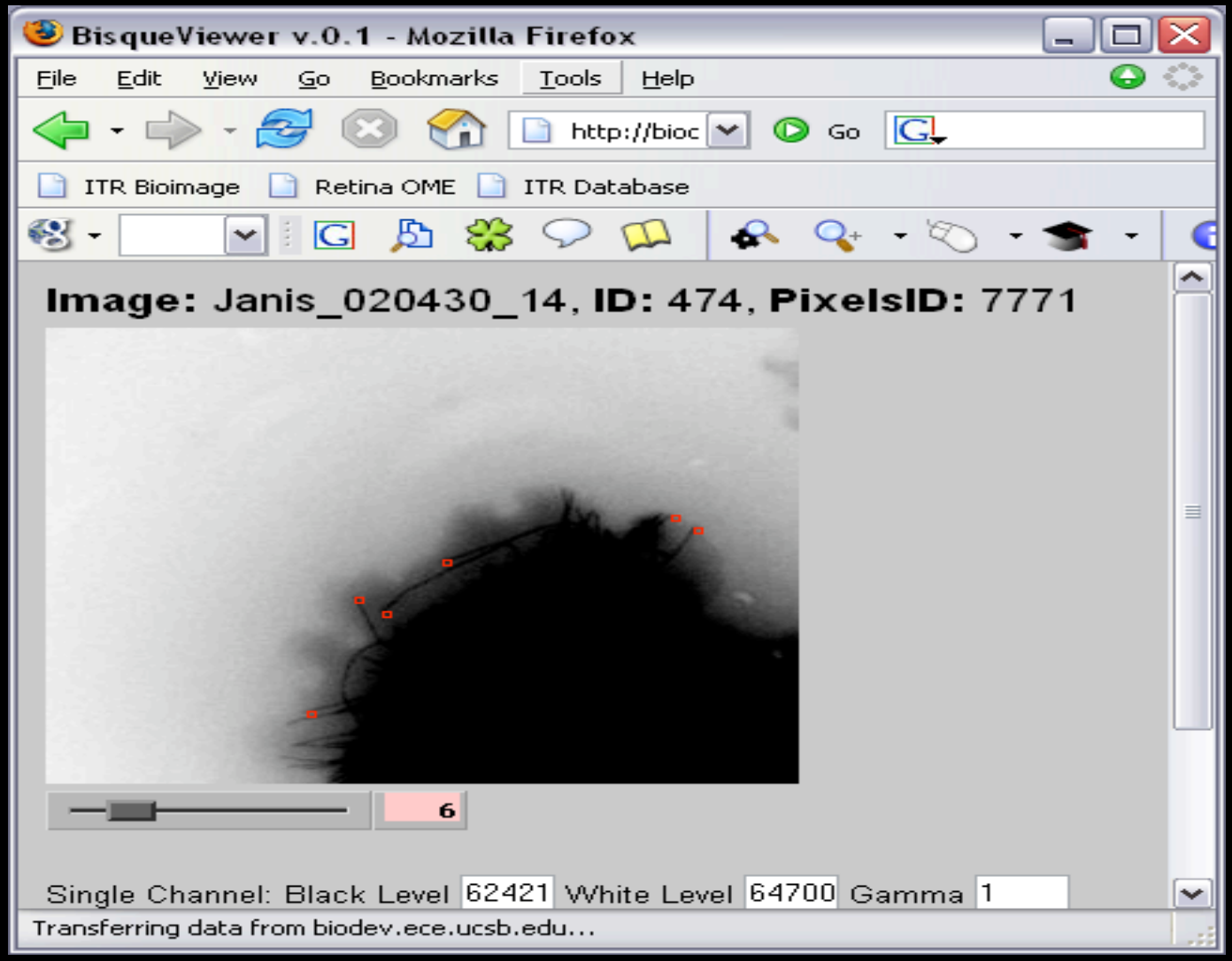
Screenshots (5D Viewer)



May 2007

Center for Bioimaging Informatics

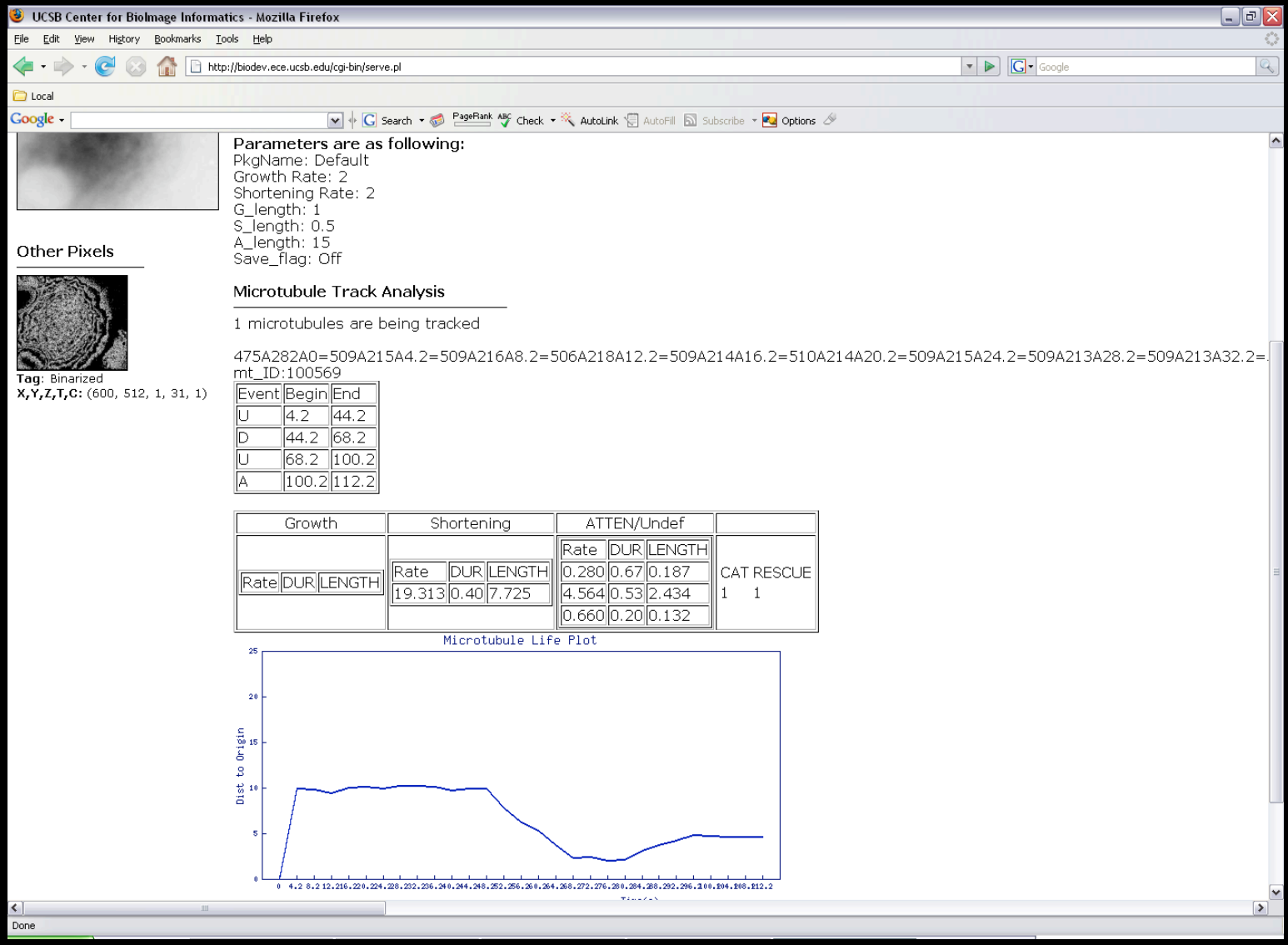
Screenshot (5D with tracks)



May 2007

Center for Bioimaging Informatics

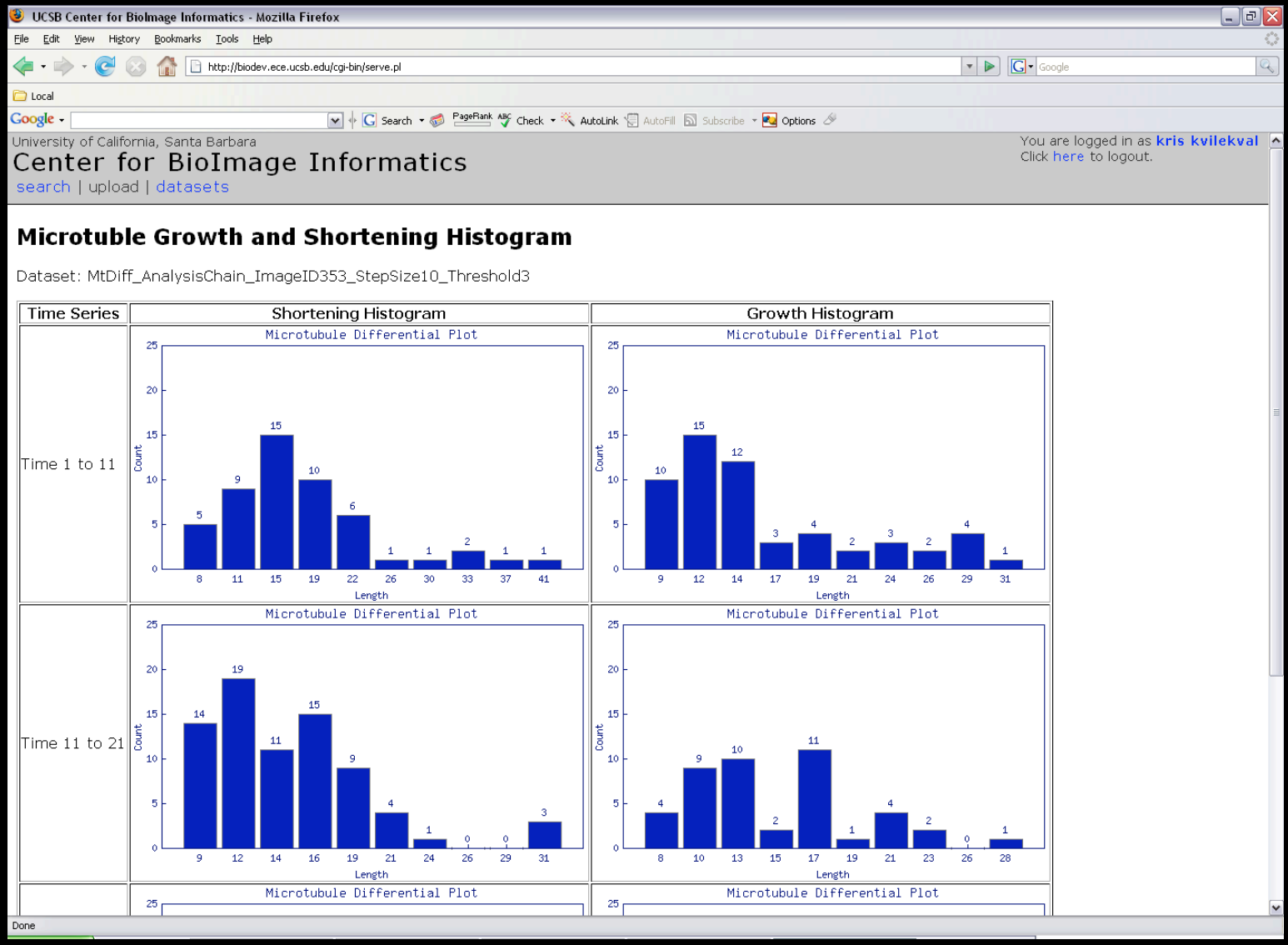
MT Life Dynamics



May 2007

Center for Bioimaging Informatics

MT Global Dynamics



May 2007

BISQUE

Bio-Image Semantic Query User Environment

- Integrated Analysis
  - Cell counter (ImageJ)
  - MT Dynamics analysis
  - MT Global Differential
- Soon to be integrated
  - MT tip tracking
  - Standalone MT growth analysis (Aleem)
  - MT body tracking (alphan)

Bisque Conclusions

- Built on OME/Postgres environment
  - Extended schema
  - Rewrote UI
  - Integrated analysis
  - Built 1st generation in use at UCSB
  - Diverse datasets
    - >5000 5D images (> 100K 2D images)
  - Integrated several *useful* analyses

Bisque/OME lessons learned

- *New data sets take significant resources to incorporate.*
  - Schema, tool, and interface modifications
- *Changing experiments also require effort.*
  - Per experiment meta data may also require significant changes
- *Building analysis should be straight forward.*
  - Researchers balk at learning complex software
  - Simpler is better

Bisque/OME Challenges

- *Automatic and semi-automatic analysis are needed:*
  - Our problems are less about bulk processing
  - High value images may need intervention
- *Semi-automatic analysis requires rich set of user tools*
  - Google, flickr, etc have raised the bar

Center for Bioimaging Informatics

Motivation:

- Current metadata model is inflexible
  - Adding new experimental images requires:
    - Changes to Digital notebook
    - Changes to Bisque interface
    - Changes to OME/postgres schema
- Shouldn't this be easier?
  - "Add my images with this experimental data"
  - "Find images tagged with rod-opsin and GFAP "
  - "Create a region and specify an object type there"
  - "Try my new segmentation algorithm given …"

New project: Bisquik

- Easily support new image collections and experimental data
- Support cross-server/lab collections/queries
- Integrate different labs metadata
- Support semi-automatic analysis
- Allow rapid prototyping of new analyses

Bisquik: Requirements

- Allow researchers to add data sets with minimal help of engineers
  - Metadata is often a (changing) list of experimental parameters
- Give analysis writers easy access to data
- Integrate data and analysis access to other systems:
  - OME
  - PSLID
- Build on previous system (BISQUE/OME)

Example analysis

- Segmentation -> automated nucleus count
  - Segment retinal layers and count
- MT Dynamics
  - Pick one more MT and measure dynamicity
- MT motility
  - Pick several MTs and measure velocity

Center for Bioimaging Informatics

Components

Web Services

Web UI

Metadata Annotation

Query

Analysis

Blob/Image Server

DoughDB

Remote Data Proxy

Analysis Engine

Flexible Metadata

- Support rapid addition of new datasets including experimental metadata
- Support new experimental protocols
- Allow analysis to create new metadata structures without a lot of work
- Extendible list of tagged values seems to be simplest model

Center for Bioimaging Informatics

Bisquik: DoughDB

OID1

Name	GH1020
Foo	2
pixels	[OID2 OID3]

OID4

image	OID1
Feature f1	[…]

OID2

image	OID1
data	Server://…
Pixel-type	raw

DoughDB requirements

- Add new tag/value pairs to any db object
  - (Foo,2)
  - (visible-cell, rod)
- Allow multiple tags with same value
  - (visible-cell, rod)
  - (visible-cell, muller)
- Support fine-grained tag permission/visibility
  - Tags have creators and access control
- Support update semantics & preserve history
  - Timestamp tags
  - No deletes (except under restricted conditions)

DoughDB queries

- Find objects (images) with tag “foo”
  - Select a.foo
- Find images with name “GV100”
  - Select a.name = “gv100”
- Find images with cellcount < 100
  - Select a.cellcount < 100
- Find images with a region similar to r1 based on feature f1
  - Select r.image = i and l2(r.f1, r1[…]) < 10

DoughDB key features

- Open ended data model
  - tag/value pairs
  - Templates for common sets
- Pair values have ts, owner, acl
- Preserves history of annotations
- SQL like query language
- Simple keyword queries
  - Antibody:rod-opsin AND antibody:gfap
  - Rod-opsin AND glial fibrillary acid protein

Bisquik ontology support

- "**ontology** is a data model that represents a domain and is used to reason about the objects in that domain and the relations between them"
- Unstructured tag/value
  - Great for taggers
  - Unhappy searchers
- Different labs use different terms for the same object.

Bisquik ontology support

- Dictionary of terms and relations
- Require (or strongly suggest) that tags and value are defined before use
- Drop into ontology editor when new values and tags are encountered.
- Integrated into search system
  - Permit (or offer) 'alias' 'part-of' 'related-to' searches

Programming Toolkit

- Goal: Allow vision researcher to easily test and incorporate new analysis
- Image/Object/Tag query/creation
- Implemented as web services
  - Resources exposed through web interfaces
  - Libraries provided for python, matlab access
- Support for data provenance
- Module execution from any environment

Module Scheduler/Engine

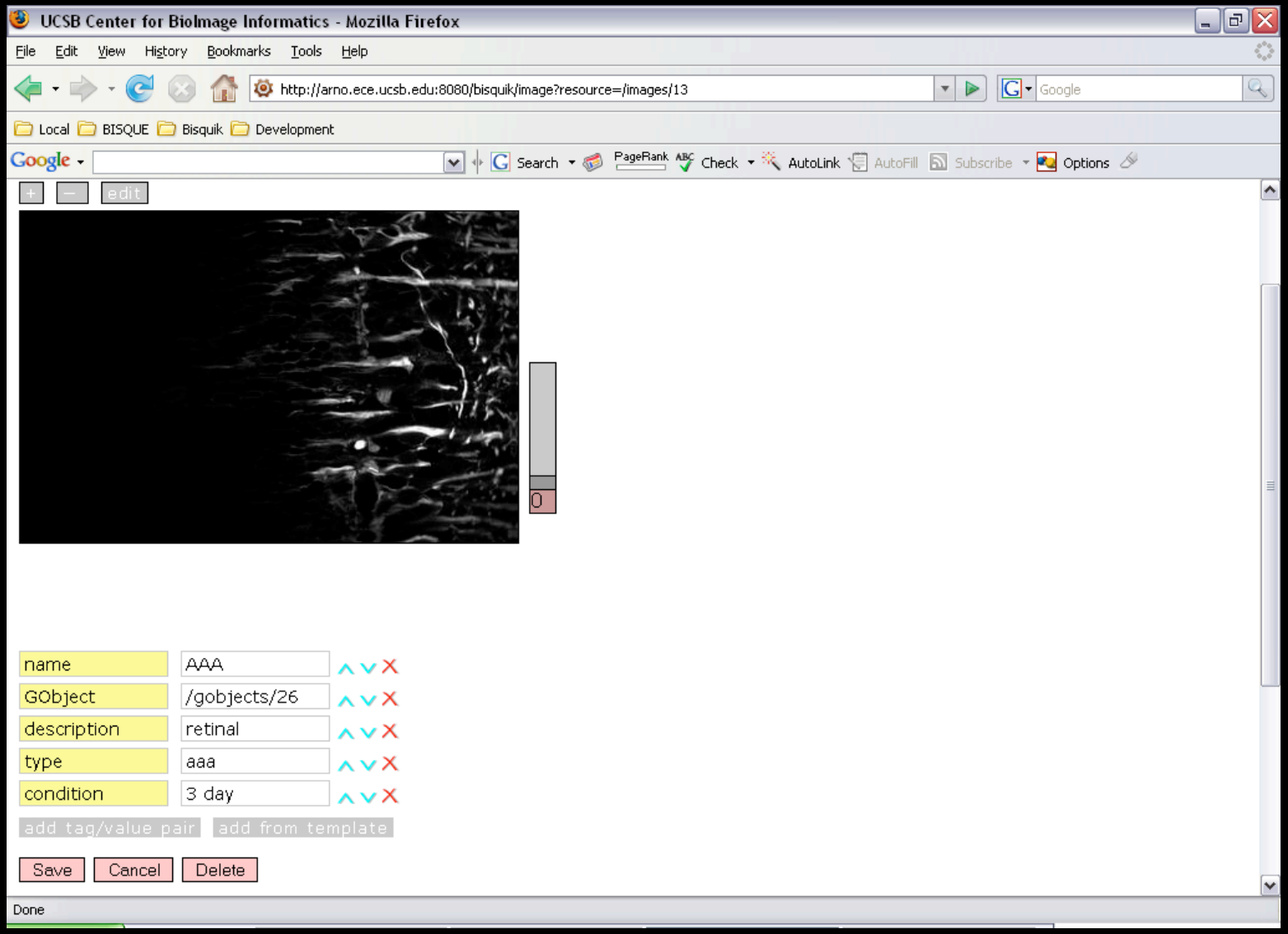
- Track free computational resources
- Execution engine
  - Schedule executions on module engines
  - Automatic component placement
- Permit development outside cluster environment
- Permit scalable deployment inside cluster environment

Bisquik interface

- Current Bisque functionality
  - Browse/Organize/Analyze
- Supports 5D images
- Flickr-like interface for image/region tagging
- Complex region definition and tagging

Center for Bioimaging Informatics

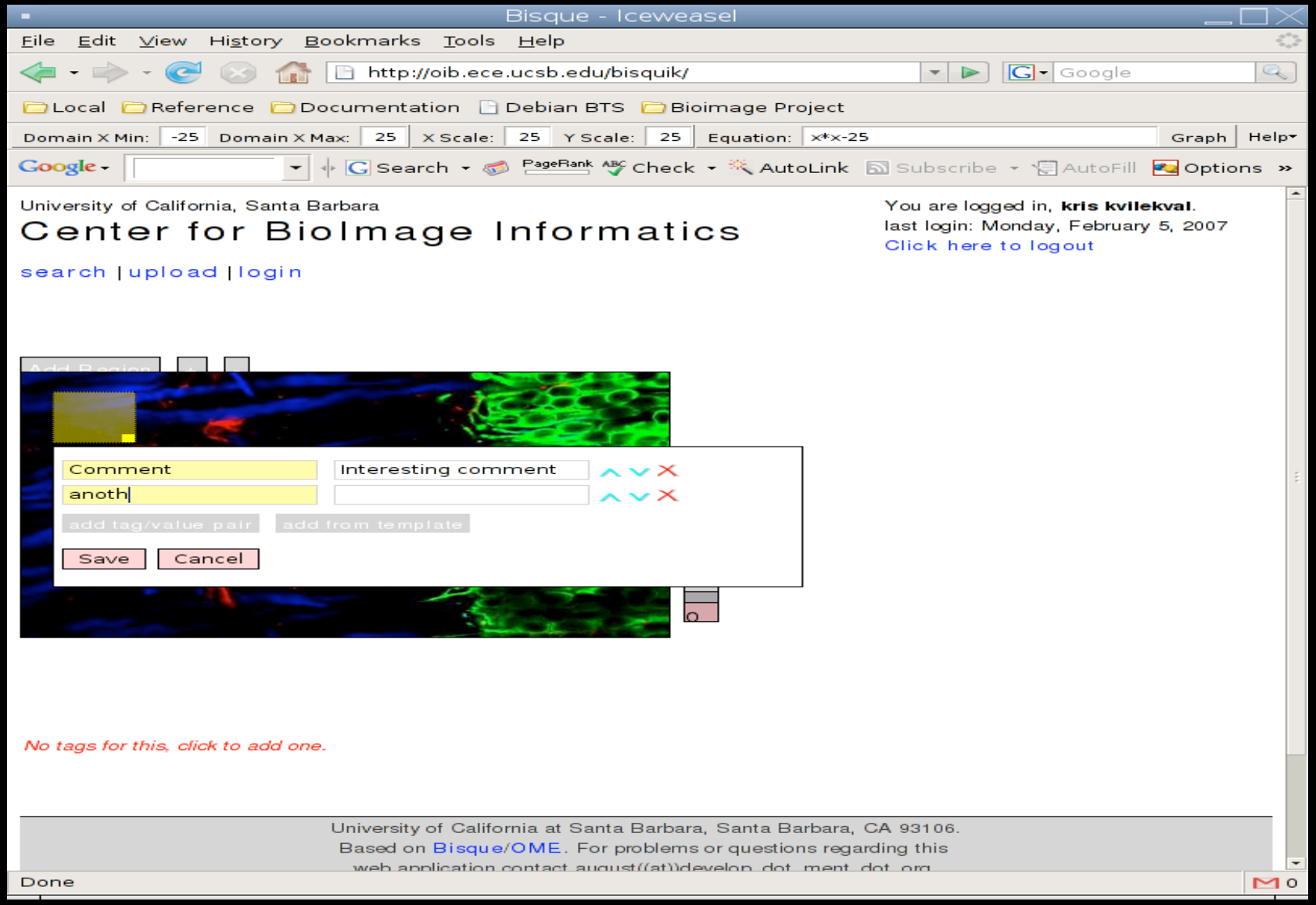
Simple annotation



May 2007

Center for Bioimaging Informatics

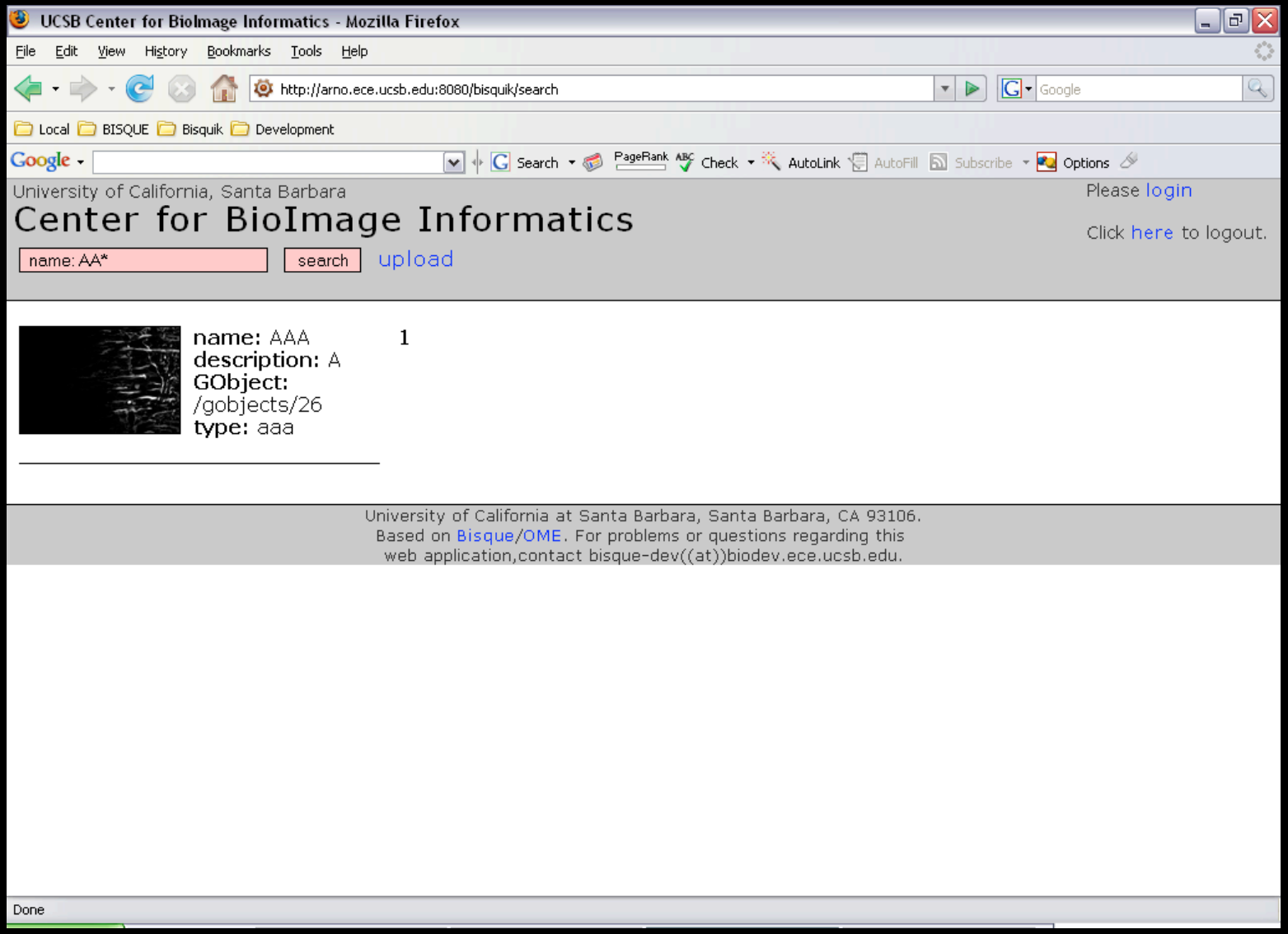
Region annotation



May 2007

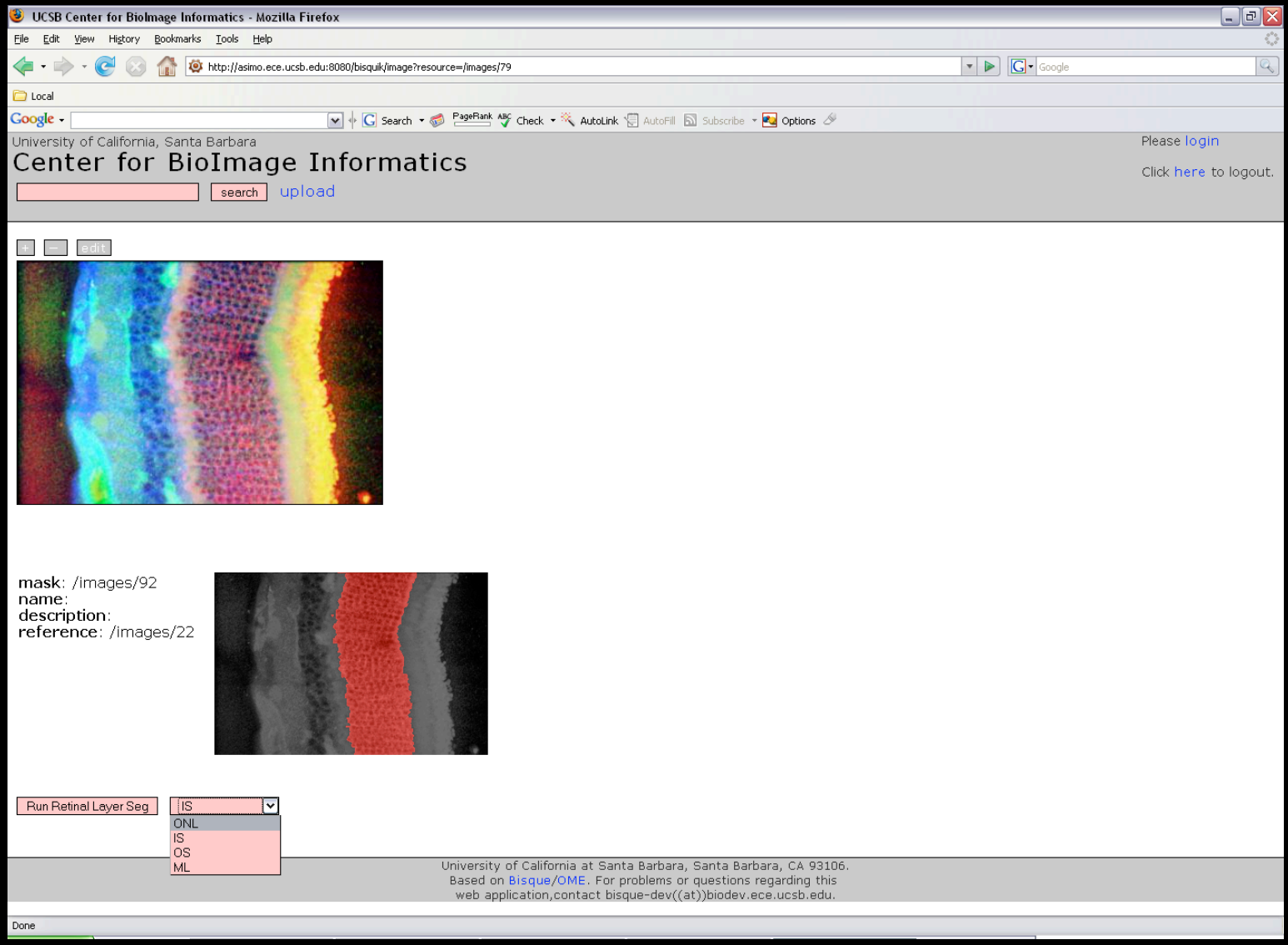
Center for Bioimaging Informatics

Search (metadata)



Center for Bioimaging Informatics

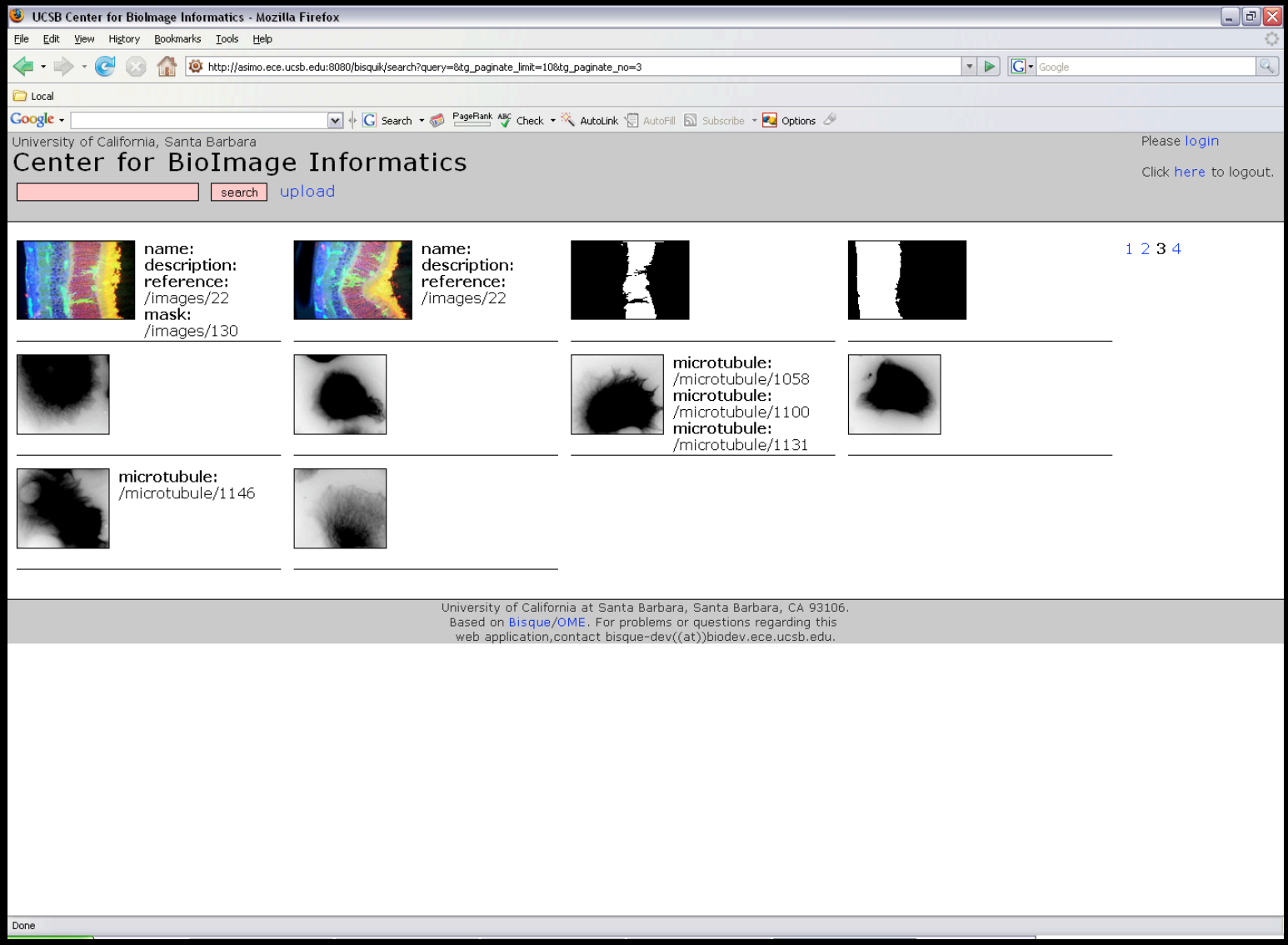
Segmentation Analysis



May 2007

Center for Bioimaging Informatics

Local and remote access



May 2007

Bisquik Metadata Annotation

- Unified offline (Digital Notebook) and online manipulation.
- Easy to build annotation forms/templates
- Allow "schema" modification "in field"
- Permit annotation templates to be shared between DN and Bisquik
- Graphical geometry annotator

Blob+Image server

- Extensible server for read-only objects
  - Pixels
  - Features
- Pluggable transforms
  - Thumbnails, slices
  - pixel transforms (watermarks)
  - Graphical metadata renderers
- Feature server

Remote Access

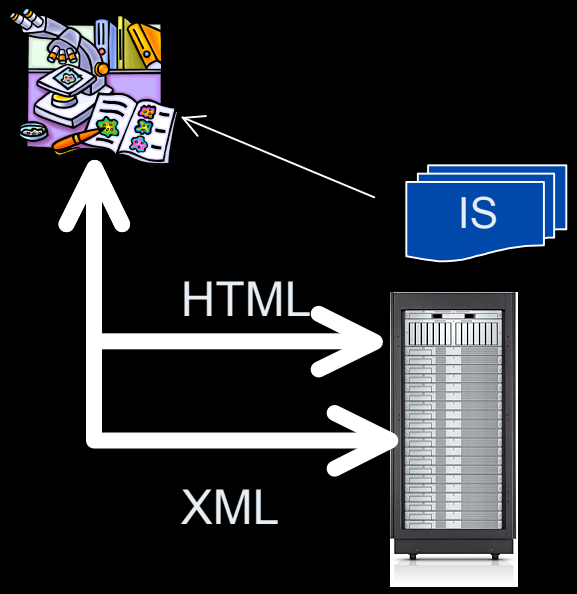
- All basic services are web accessible:
  - RESTful (simple model, web caching)
  - DoughDB, Image server, Module engines
- Cluster Database support
  - Image collections are split across machines
  - Unified view
  - Query engine distributes and resolves
- Already supports access to 'foreign' data sources:
  - multiple BISQUE/PSLID/OME sources

Scalable/distributed deployment

- Components services
  - Image Server
    - Manipulations (slice, format, etc)
  - Data Service (query, storage)
  - Module Engine (Analysis Executions)
  - Web Service (browser, aggregation support)

Center for Bioimaging Informatics

Simple deployment



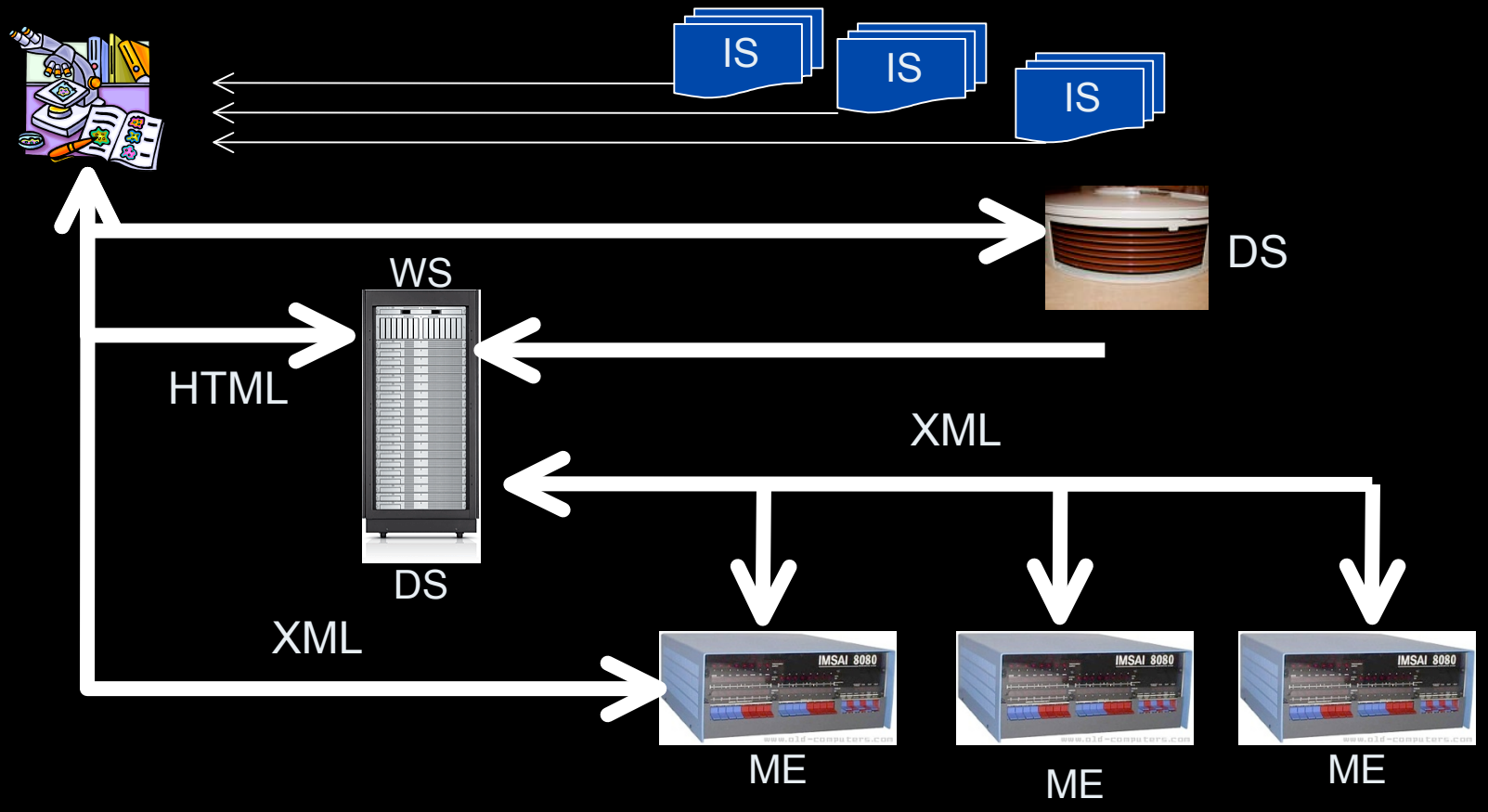
Web Service
Image Server
Data Server
Module Engine

- Image Service
- (Meta) Data Service
- HTML Server
- Module Engine

RC – Rich Client WS-web server IS - Image Server DS – Data Server
MS - Module Scheduler ME – Module Engine

Center for Bioimaging Informatics

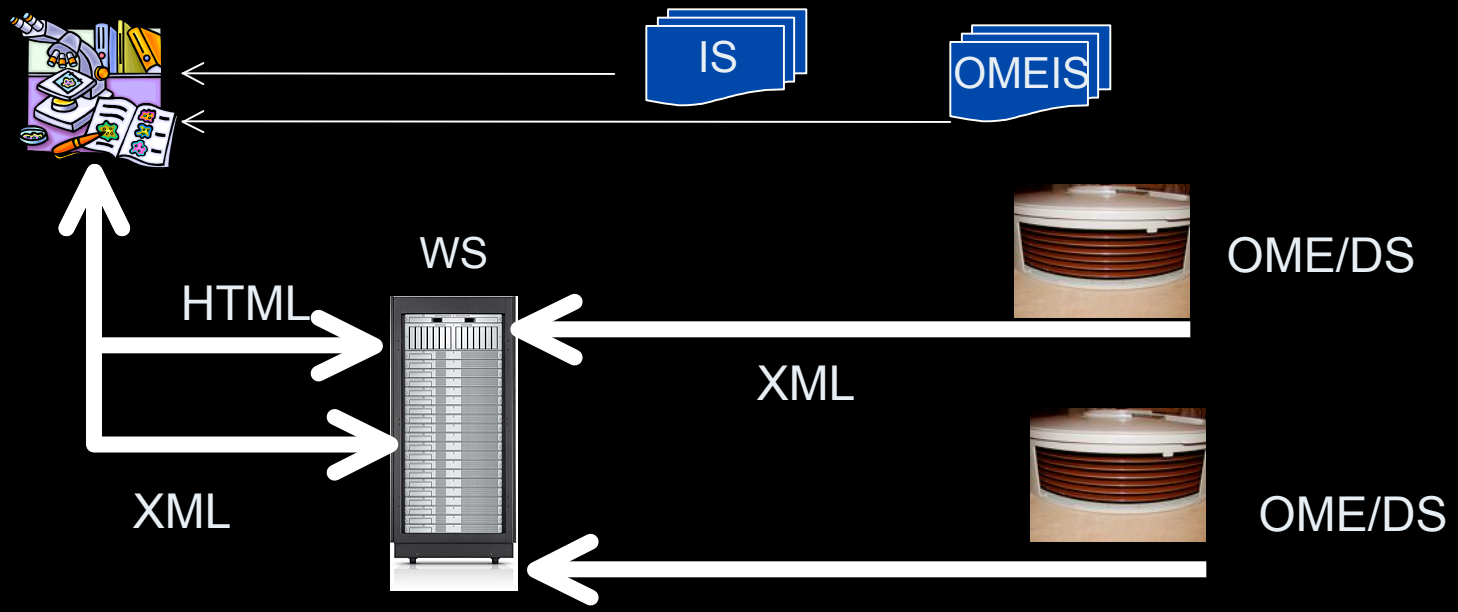
Scalable services



RC – Rich Client WS-web server IS - Image Server DS – Data Server
MS - Module Scheduler ME – Module Engine

Center for Bioimaging Informatics

Remote services



RC – Rich Client WS-web server IS - Image Server DS – Data Server
MS - Module Scheduler ME – Module Engine

Service Examples

- http://host/images

```
<response>
 <image resource="/images/1/" imgurl="/imgsrv/2" />
 <image resource="/images/2/" imgurl="/imgsrv/3" />
</response>
```

- http://host/images/1?view=full

```
<response>
 <image resource="/images/1/" x="512 y="512"
     imgurl="/imgsrv/2">
    <tag resource="/tags/10"
         name="description" value="mt image" />
 </image>
</response>
```

Service Examples

- GET http://host/modules

```
<response>
 <image resource="/images/1/" imgurl="/imgsrv/2" />
 <image resource="/images/2/" imgurl="/imgsrv/3" />
</response>
```

- POST http://host/modules/1

```
<request>
  <tag name="p1" value="/images/1" />
  <tag name="count" value="0" />
</request>

<response>
 <image resource="/images/1/" />
 <microtubule resource="/microtubule/1" />
</response>
```

Bisquik Status:

- Web UI
  - Uploading, Tagging, simple searches
  - Demo at http://biodev.ece.ucsb.edu:8080/bisquik
- DoughDB
  - Prototype based on SQL/BerkelyDB
  - Multi node storage aggregation and queries
- Analysis
  - Layer segmentation
- In Development
  - Ontology support, advanced query/indexing

Bisquik: Initial impressions

- Focus : Ease of Use
  - For biologists: simple data model, easy searching
  - For analysis developers: develop in comfortable environment
- Web UI
  - Tools developed for semi-automatic analysis
- DoughDB
  - Performance needs to be tested on large sets
- Analysis ***
  - < 1 day for researcher to use tool kit
  - ~1 day for interface improvements
- Development
  - Rapid development tools
  - Agile language and methods (python)
  - Lots of progress in little time (march 15-Now)

Conclusion/Vision

- Prototype data model and analysis in Bisquik
- Use data and analysis from multiple sources
- Migrate to backend systems (OME/PSLID) as needed

Bisquik plan

- Release 0.1 : April 2007
  - Bisquik Tagging + DoughDB
  - Bisque/OME Bridge (image + metadata)
  - Simple queries (antibody:vimentin and 'cross section')
  - Blob server
- Release 0.2 : May 2007
  - Web + DN Metadata annotations (text,graphical)
  - Distributed queries
  - Several analyses (segmentation, MT motility)
- Release 0.3 : June 2007
  - Ontology support (DoughDB + UI + query support)
  - Access to 'foreign' analysis BISQUE/PSLID
  - Other analysis from local researchers
- Release 0.4 :
  - Analysis engine scheduling + performance tests

Info

- Open Source Code
  - http://biodev.ece.ucsb.edu/projects/bioimage
- Demo and UCSB collections
  - http://biodev.ece.ucsb.edu/bisque
  - http://biodev.ece.ucsb.edu:8080/bisquik

Development Team

August Black, Jae Choi, Dmitry Fedorov, Brian Ruttenberg