CID Imanage: Image Database on a microscopy facility

P. Paul-Gilloteaux [1], N. Gupta [2], S. Goud [3], A. Kulkarni[2], A.Ghosh [2], T.Lecorgne [6], O. Renaud [1], F.Cordelieres [4], J. Boulanger [5], P.Chabosseau [4], O. Leroy [1], P. Le Baccon [1], V. Fraisier [1],F. Waharte [1], T. Piolot [1], L. Sengmanivong [1], C. Barette [3], J. Salamero [1,5].

[1] PICT-IBISA CNRS/ Institut Curie - Paris (France), [2] Strand Scientific Intelligence - Bengalore (India), [3] IT department, Institut Curie - Paris (France), [4] PICT-IBISA CNRS/ Institut Curie - Orsay (France), [5]Serpico Team CNRS/ Institut Curie - Paris (France), [6] Serpico Team, INRIA Rennes (France)

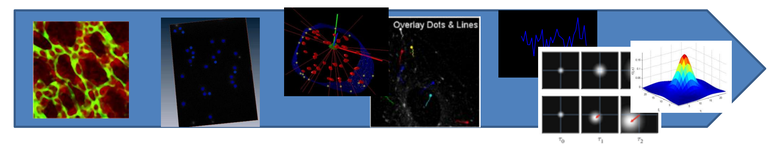
Why did we need it?

→ a production of images over 4 years from 400 to 500 TB to archive, with an annual average of 100TB to store redundantly

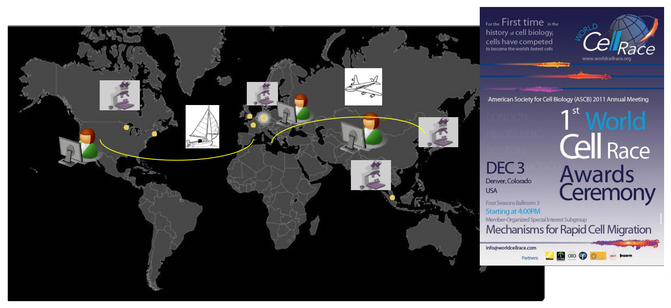
→ produced by ~40 microscopes (photonic and electronic) with different formats,

→ around 250 users in Curie and external, with a turn over.

→ Projects from tissue, developmental biology, single cells studies, cellular and subcellar dynamics, molecular dynamics



Example before it was set up:



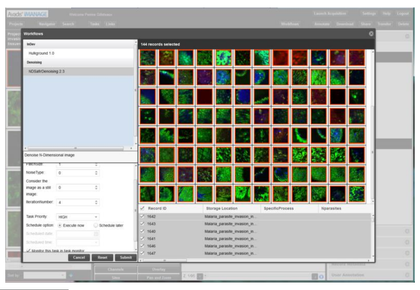
→ People would sent their data either by multiple file transfer, or by mailing external hard disk.
→Need for conversion and normalization of data before processing
→Accurate excel file with the cell categories and condition of acquisition would need to be filled and referenced to datasets.

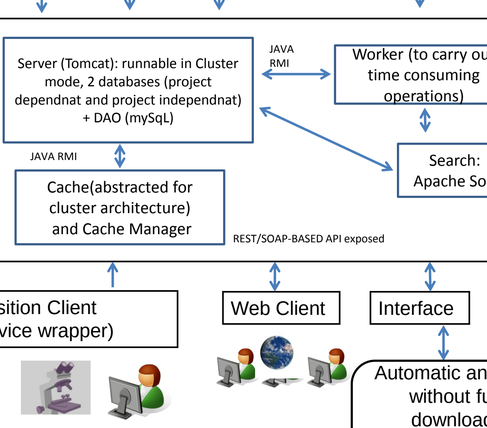
Centralized repository of data, metadata and results of analysis.

Database organisation:

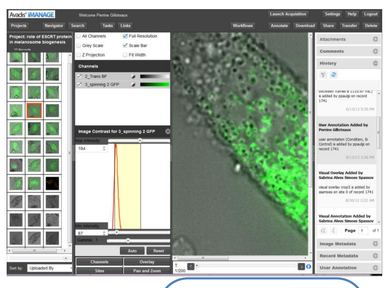
→Storage is scalable
→Server can be replicated, without duplicating the data
→Workflow algorithm can be set up on cluser
→Data can be checked for quality before uploading them to the server.

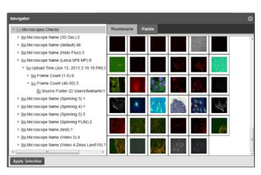
→Microscope upload masks can be created.

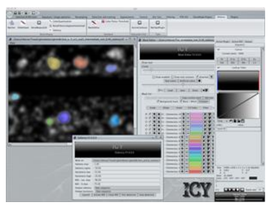




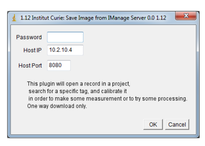
Processing batch of images on cluster (denoising, deconvolution, Tomographic reconstruction…)







Metadata (pixel size, acquisition time,…) annotations, Parsing nD images Insuring reproducibility by storing all processing



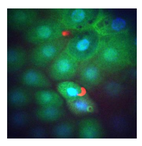
Example of use cases

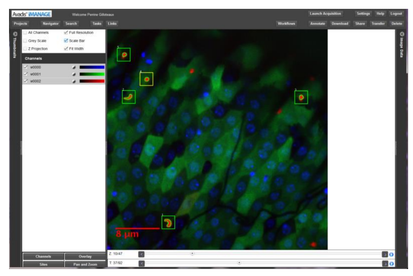
Use Case 1: External user coming for 2 weeks on the facility (Proof of Concept EuroBioimaging). M. Anitei, Biotech, TU Dresden, Germany

→ Different microscopes used with different formats : user can visualize her data and batch download them as tif files.
→some process done after the user came back to its lab (as for example beads based registration of dual view camera, Results are linked to the original images and also accessible via the webclient.
(Image J TurboReg script accessing beads for one date, computing the deformation, and splitting and correcting corresponding images before uploading them back in the database)

Use Case 2: External collaboration: G.Volohonsky IBMC Anopheles group Unistra, Elena Levashina Max Planck Institute for Infection Biology

- Proteins attacking the parasite in mosquito LRIM, APL1, TEP 1 (hemocyte)
- Species of mosquito: G12 (india),DSX and HYPER (African)
- Marker of damaged cells: Sugar Dextran, or hyper protein becoming fluorescent when binded to H202 or whith OPH changed.
- Different time of infection, different times post infection.



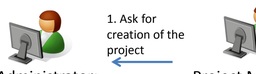


Images are processed, to 3D segment parasites and cells, and ImageJ Rois are converted to database visual overlay for visual check.

Goal: constructing an averaged model of the behavior of malaria parasite in mosquito: Processing based on annotations
Questions to answer:
→Parasite going out the gut wall: rate and proportion against different time?
→ Is there any shape factor of the mosquito as an additional parameter?
→Inside the midgut: which are the mode of displacement of the parasite among cells?

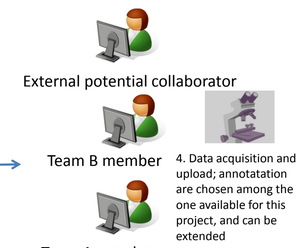
Use Case 3: Internal users of the facility: collaborative project between Team A and Team B

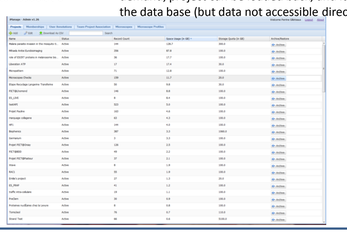
Storage is invoiced by the IT department to team.
Amount used by each team can be listed from the webadmin.
Team leader can check the usage of the storage of its team at any moment from the webclient

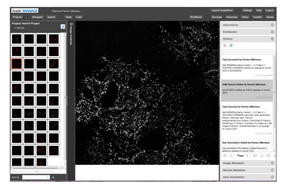


1. Ask for creation of the project
2. Creating account, Project / Storage quota association
3. Add users to this project, with different rights.
4. Data acquisition and upload; annotatation are chosen among the one available for this project, and can be extended
5. Process data, on cluster for some identified routine algorithm.
6. Project has led to publication, flag the project for archival (long term band storage), freeing the short term storage. On demand, project can be loaded back, and it still searchable in the data base (but data not accessible directly)

Send a download link for some sample data. If interested, it will get an account for the database and will get associated to this project.

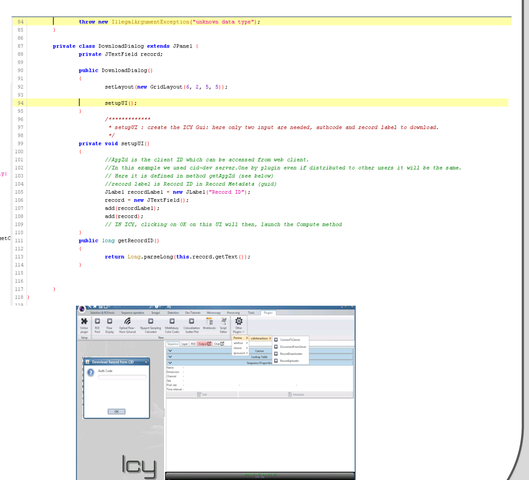






Exemple of cluster processing. History is stored.

Example of Third-Party software interaction: ICY: connect to the database, access record, work on them, upload them back.



Exampel of ICY plugin dowloading an image in the database, and using metadata to calibrate it.