Minutes of the OMERO Workshop

5-6 September 2019, ITQB, Lisbonne

Participants

- Petr Walczysko, Michal Franek, Valya Vassileva, Ana Paula Santos, Christophe Tatout, Pierre Pouchin, George Tsouloupas, Nadia Goué and Gabriel Martins
- Visio conf: Hank Bass, Stefanie Rosa, Célia Baroux and Sophie Desset

Welcome speech by Ana Paula (Local organiser at ITQB) + Christophe introduction about INDEPTH:

COST action INDEPTH dec 2017 > dec 2021: Understanding the nuclear domains and their imapact on gene expression. WG1 3D bio-imaging

Why this workshop: Get an open-access depository for INDEPTH datasets (images, sequences) & Collect plant datasets in 3D imaging of the nucleus (WG1)

Day 1 : Training with Petr and Pierre

 General introduction about the OME project, OME platform

- Trainings: https://www.openmicroscopy.org/training/
- Video tutorials: please visit: https://www.youtube.com/channel/UCyySB9ZzNi8aBGYqcxSrauQ

Tools:

- Omero insight, best for image upload: https://www.openmicroscopy.org/omero/downloads/
- Omero web, best for image manipulation
- Omero plugin for fiji, for producing modified images and make analysis https://imagej.net/Fiji/Downloads https://github.com/ome/omero-insight/releases/download/v5.5.5/OMERO.imagej-5.5.5.zip

Functionalities:

- *Omero.parade* is a new functionality to link output results and linked them to images and then to produce graphs
- *Omero.figure* to produce your figures for publication, share a file through a weblink

It was an intensive day but everybody got excited by the Omero functionalities! Thanks Petr and Pierre for this amaizing training session.

Day2: open discussion all together

What are the INDEPTH objectives?

At least 16 people / labs expressed interest into sharing datasets; A survey is available at
https://docs.google.com/forms/d/1Q_D2FMp5bcDjhbBTi3S_46rCkBI508R3LvTCZOZZnfc/edit#responses

- Give members space to upload, store and share data between groups within INDEPTH
- Produce "public folders" corresponding to published datasets

INDEPTH project may have several types of images to store:

- few images illustrating a publication; DOI available
- A complete dataset to be allowed to benchmark a new software in the future
- Datasets to be exchanged between the 2 teams before publication

First, we clarified the aims of OMERO and IDR, two software used as 3D image repository

IDR is curated by the OME team, OMERO is not and the **curation** should be done by someone from INDEPTH. IDR is fine as a 3D image repository. IDR is a kind of Genbank for images. No limit of size to be stored for most file types (there are limits noted, but they identify a cutoff at which they will consult w/ submitter).

This link describes IDR submission & mission:
https://idr.openmicroscopy.org/about/submission.html

IDR pub announcement in Nat Meth, June 2017 https://www.nature.com/articles/nmeth.4326

Description of IDR from the "SUBMISSION" page of IDR

We aim to publish reference image datasets, which have value beyond simply supporting an original publication according to the **Euro-BioImaging - Elixir Image Data Strategy**. This includes:
- Datasets **associated** with an existing or upcoming publication
- **Complete** datasets - not just images supporting one figure in the publication
- Datasets whose metadata can be **integrated** with other datasets via identifiers from well-known biomolecular resources (Ensembl, NCBI Entrez Gene, RefSeq, PubChem, ChEBI etc)
- Datasets generated using new imaging **methods** or new analysis methods
- Datasets that are likely to be **re-analysed or incorporated** into other studies or integrated with other imaging datasets

To make IDR datasets as widely re-usable as possible, we strongly recommend that submitters make their datasets available under **CC-BY** license. The licensing information should be included as part of the metadata. Image datasets that don't meet the criteria for reference images may still be published on **BioStudies** or **Dryad**. Contact these projects for more information.

Comments from the Workshop members: IDR is a Long term lasting repository. IDR is a web instance with omero running in the background. Good for publishers (provide DOI); rules are defined for annotation; much be linked to a publication of most image types i.e. Electron microscopy, Light Sheet. ROIs (like shapes, distances…) can be included. Note that some publishers start to provide the opportunity to produce 3D pdf for figures.

===
HB suggests to adopt the following language to avoid confusion.

"**IDR**" is self-named repository, the fact that it runs omero in the background is somewhat not relevant, but a bonus from our perspective because the INDEPTH group will primed to use, trained & able to navigate, & add to IDR.

"**OMERO**" alone can be use for general reference the platform, software, system, ...

"**INDEPTH-omero**" is for the within group omero, currently hosted at FSU, but easy to install elsewhere, so INDEPTH-omero is the one hosted at omero.bio.fsu.edu, & synonymous with omero-FSU or FSU-omero.

"**FSU-omero**" server is multi-purpose, one of which is INDEPTH.

If we do make other server locations, we can simply extend the names, e.g. I**NDEPTH-omero@fsu** & I**NDEPTH-omero@oxfordbrookes**, if needed or **INDEPTH-omero1**, **INDEPTH-omero2**, etc (as these will not be linked, & could house unique or mirrored datasets. For now w/ only one group omero INDEPTH-omero & INDEPTH-omero@fsu/INDEPTH-omero1 are the same.

===

HB proposition for INDEPTH-omero@fsu:

Space available at INDEPTH-omero@fsu = 4 Tb available

Here's a summary take w/ regard to omero & image storage and sharing.

- We will (as planned) develop for INDEPTH group a shared storage and collaborative image bank, hosted at omero-fsu for now, noting that it is relatively easy to set up servers, omero tech & spec docs & support are thorough and excellent. Most computer sys admins can do the install & upgrades & the user admin control panel (that part HWB does) makes adding groups & users very easy through a GUI.

- Long term storage will use IDR and INDEPTH-omero training will enable trainees as knowledgeable early adopters of the 'public image data bank submission' concept. IDR was deemed ideal for many of our needs as they have staffed curation and data standardization suitable for grant-required data management plans, journal requirements for pre or post publication data access. Given that the IDR runs

OMERO in the background, our INDEPTH-omero training and use represents an early adoption and contribution to the GenBank-like image bank concept.

● The expectation of depositing research data to image repositories will likely become standard scientific practice. Long term, the IDR, or a comparable staffed and dedicated image databanks, will be the destiny of submitted and published data from. The group INDEPTH-omero will enable and accelerate data sharing, should stimulate new collaborations, and familiarize us with a clear path to image data use, reuse, & longevity.

==
In summary

To be trained on OMERO is useful for INDEPTH members and useful for IDR submission OMERO-server are useful to store and organise images at the lab/Institute level and to share data between groups like INDEPTH groups while IDR will offer a Long term lasting repository like Genebank for sequences accessible by the whole research community

INDEPTH-omero is needed for the INDEPTH network

==

More details about how to set-up the INDEPTH-omero server

The way the current INDEPTH-omero@fsu is organized is not optimal.

Petr made some suggestions t to prevent the instant creation of useless accounts (that can't be deleted). Once the system is clear for the participants, more accounts can be created as the INDEPTH organizers see fit, in the non-public groups. For the public groups, the restriction on amount of data owners is advised, to keep clarity.

The best configuration would be the *OMERO Public access* for the OMERO server + groups are then made public (to share publication) or non-public (to share datasets within an institute or between institute(s)). This should be by the OMERO administrator.Then make a *lab accounts* for each institute for instance to share unpublished data. 1 or 2 accounts by institute. Lab account but one can modify another file. That is a risk to take on shared data.

OMERO read only means that if 2 users are within the same group they can see each other data but can modify only their own datasets

==

Notes from HB:

Private (rw----): All data in this group is only visible to the user who owns it and the group owner. Other users cannot see who else is in the group or view their data. The group owner can view the data for other group members but not make any edits (same as read-only behaviour).

Read-Only (rwr---): Users in groups with this permission setting can view each others' data, but cannot edit or annotate another user's data. You can view other users' images but not comment, rate or tag their images.

Read-Annotate (rwra--): Users in a Read-Annotate group can view and annotate the data belonging to other users. You can tag another user's images or use their tags to annotate your own images. You can add comments to their images and save your own rendering settings for each image. However, you cannot edit the names of their images, projects, datasets or tags etc.

Read-Write (rwrw--): Users in a Read-Write group have permission to edit and delete the data belonging to other users.

==

PW : The images should not be public immediately. A like of ante-chamber before public (a non-public group where you prepare and check the data before publication). Be Careful because of too much heterogeneity in the public area (groups where the public user is member of).

There are the user(s) who owns the data and there are the public viewer (public user) who views the data in a public group. The public user does not own any data him/herself.

An option would be to have 2 OMERO web servers and one would be public 2 would be sharing the same OMERO server - this option is described in
https://docs.openmicroscopy.org/omero/5.5.0/sysadmins/public.html

===

To Do list:

Participate in follow up discussion to
- define how to **structure INDEPTH-omero**,
- how to participate, default permissions,
- conduct code for sharing unpublished data (new ground here, actually, so not trivial step)
- pipeline or guideline to public sharing of data associated with peer-reviewed articles in IDR

Decide on group and account structures as per Petr best practices recommendations
- HWB will
- install version 5.5 at omero.bio.fsu.edu
- enable the omero public option which allows specialized user named "public" (different from the regular user account we created for access to published images) to be added to groups, making their content accessible to all online.

Grant for teaching was at the origin of Omero@fsu:
- **make a folder for training students**
- **Training set of data:** another folder

Nadia Goué: Identity federation (to be defined) to give access the server to members outside the university under collaborative regulation like INDEPTH project

Define a **Chart of good practices** and confidentiality before accessing the INDEPTH datasets

Define tags and/or key:value pairs used at INDEPTH-omero.We should define a few rules for using the facility. Some kind of DOI should acknowledge the owner of a given dataset (See IDR rules for DOI)

Define a how to prevent downloading datasets from OMERO without asking the **permission to the owner**? Some by-pass solutions are available as suggested by Petr
- https://docs.openmicroscopy.org/omero/5.5.0/sysadmins/public.html
- https://omero.lifesci.dundee.ac.uk/pub/schleicher-et-al-2017

Célia/Biplane: Explore the question on post-processes / IMARIS

All: organize training session at our institution together with the OMERO team. The only requirement is to fund the travel + accommodation for the trainer(s), training is free

Christophe: inform about INDEPTH-omero@fsu all MC members at the next MC in El Escorial in December 2019
===