

About open science, **cytominE** & *flows*

uliege.cytomine.org

biaflows.neubias.org

 [cytomine_uleige](#)

November 2020

Raphael.Maree@uliege.be

Scientific coordinator @ University of Liège, Belgium (since 2010)

Co-founder of Cytomine not-for-profit cooperative (since 2017)



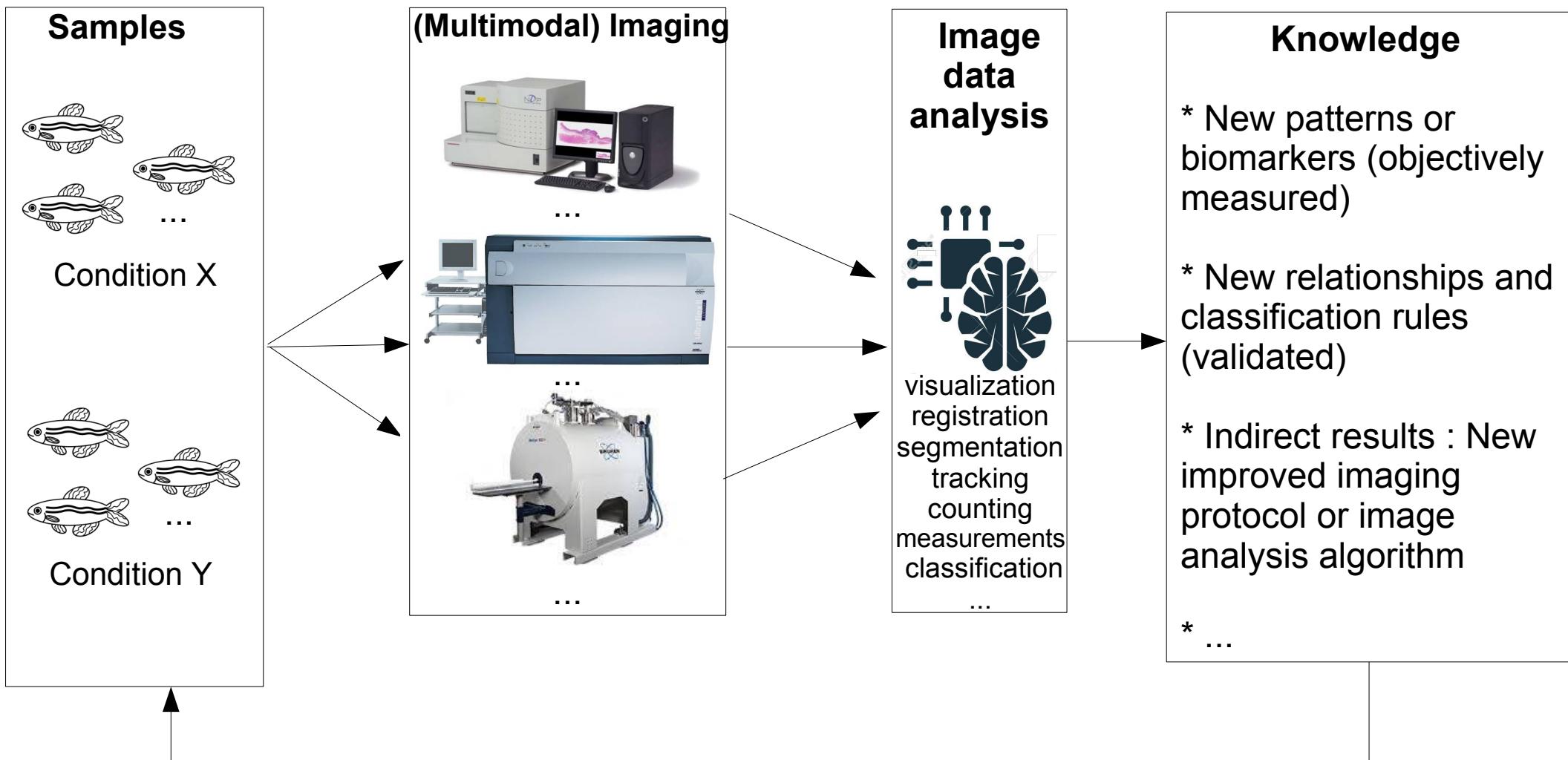
MONTEFIORE INSTITUTE

Department of Electrical
Engineering and Computer Science

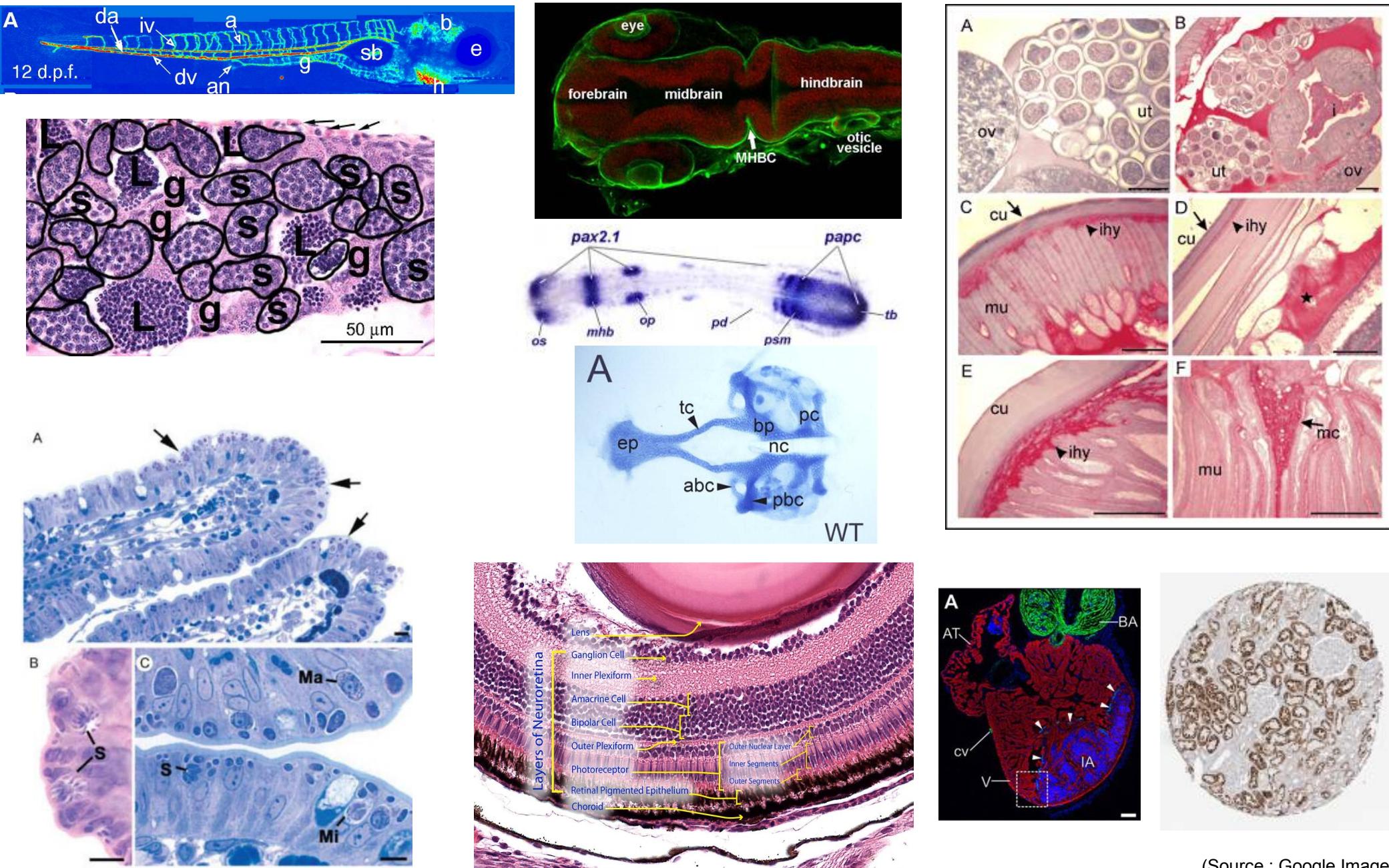


Life science research heavily relies on images

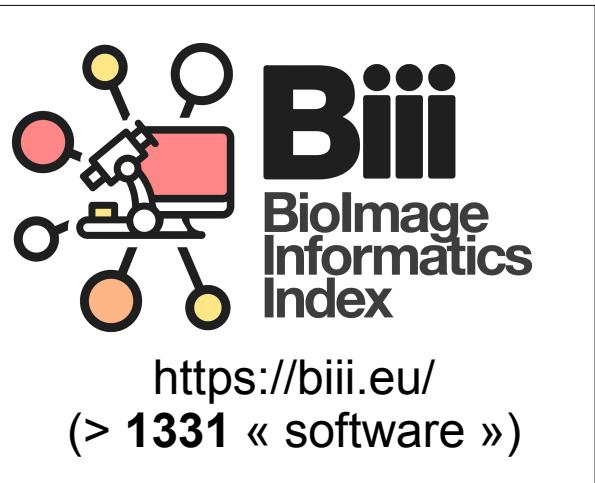
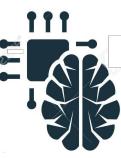
(Exploratory analysis / Biomarker discovery / Classification of samples)



A multiplicity of image analysis tasks



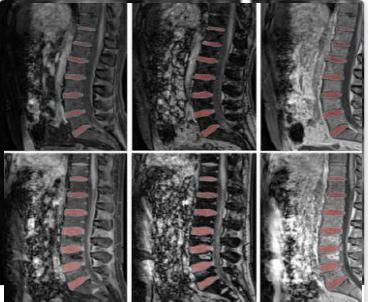
A multiplicity of image analysis tools



Bii
Biomage
Informatics
Index

<https://biii.eu/>
(> **1331** « software »)

openneuroscience
<https://open-neuroscience.com>



(e.g. intervertebral disc localization)

Medical Image Segmentation

57 benchmarks

113 papers with code

<https://paperswithcode.com/>
(> **11000** ML papers with code)



GitHub

<https://github.com/search?q=zebrafish>
(e.g. > **400** zebrafish code repositories)

...

Image analysis practices are not ideal (1/2)

Image analysis is still often :

- Qualitative (seeing is believing?) or performed (semi-)manually
- Performed within tissue subregions or in small sample groups
(might not be statistically significant or one might miss specific patterns)
- Performed by isolated experts and stored locally
(sometimes lost)
- Not saved, or in closed formats so hardly reusable
(e.g. Photoshop draws for paper figures)

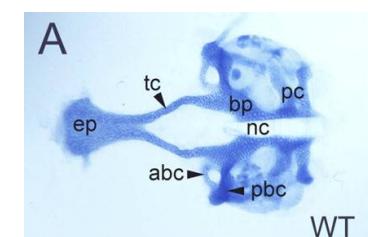


Image analysis practices are not ideal (2/2)

Challenges for biologists:

« Not easy to reuse previous work on my images because ... »
... many nice biology papers do not come with a tool or few details
... which algorithm to choose in the zoo ? How can I trust it ?
... best algorithms might not be implemented in user-friendly software
... many algorithms are not interpretable (black box)
... publish or perish ... (a lot of work is lost after publication e.g. data, image analysis results & tuned methods)...

Challenges for computer scientists :

« Not easy to design generic methodologies because ... »
... so many imaging techniques & combinations
... few annotated (ground-truth) datasets available (proprietary formats)
... meaningful results ? expert's proofreading ?
... publish & perish... validation only on 1-2 (small) dataset(s) : ad hoc

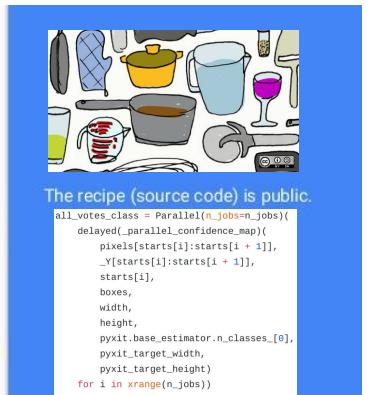
→ **suboptimal science practices** (reproducibility issues, waste of time, ...)

Can we solve this by tomorrow? No, but...

Let's try to **collaborate** more effectively towards **more generic** and **reproducible** image analysis workflows and biomedicine research.

Practical suggestions to improve our situation :

- **Image sharing**
- **Image annotation (ground truths) sharing**
- **Image analysis workflows sharing and benchmarking**
- **Open science** (open data access, open source code, ...)



→ Freedom to install, inspect, extend, improve, reproduce, redistribute

(see also Brito et al., Gigascience 2020)

Existing image sharing solutions

Centralized repositories



www.ebi.ac.uk/bioimage-archive/



www.oasis-brains.org/



bossdb.org/projects



www.cancerimagingarchive.net/

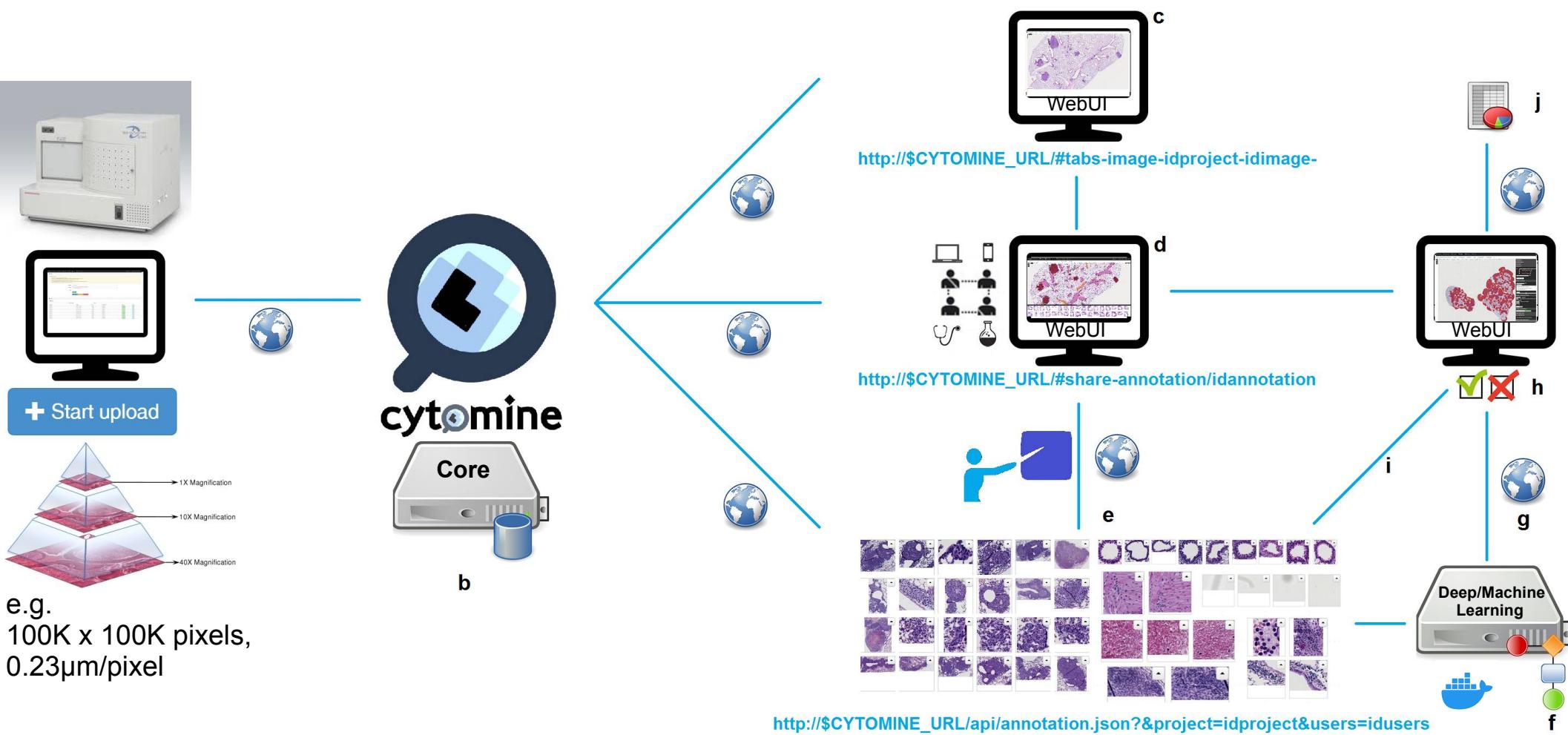
.... Beyond image sharing ?



Overview of main features

[cytome.cytomine.org \(official version\)](https://cytome.cytomine.org)
[uliege.cytomine.org \(ULiège R&D version\)](https://uliege.cytomine.org)

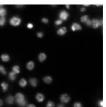
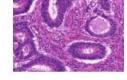
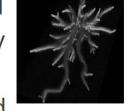
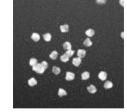
cytominne enables collaboration through the web (Sharing of images, annotations, algorithms, results)



cytominE: organize your images, securely, on the web

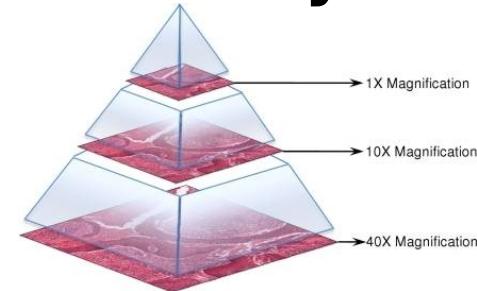
Create and manage multiple **projects**:

- Support **various digital pathology, microscopy & other image formats**
- Users with **authentification** (e.g. LDAP), **access rights**, and **roles**

Name ↑	Description	Members ↓	Images ↓
➤ DATA-SCIENCE-BOWL-2018	Heterogeneous collection of 2D images used to illustrate nuclei segmentation. Includes stage1_test image set from BBC038v1 , available from the Broad Bioimage Benchmark Collection [Ljosa et al., Nature Methods, 2012].	 	5 130
➤ GLAND-SEGMENTATION-TEST	The aim of this problem is to classify pixels belonging to glands in histopathology images cropped out from images of 2015 MICCAI challenge of gland segmentation (GLaS 2015).		5 480
➤ GLAND-SEGMENTATION-TRAIN	The aim of this problem is to classify pixels belonging to glands in histopathology images cropped out from images of 2015 MICCAI challenge of gland segmentation (GLaS 2015). These images were used to train the machine learning classifiers available in GLAND-SEGMENTATION-TEST.		5 671
➤ LANDMARKS-DROSO	Landmark detection in Drosophila wings, data from UPMC (Vandaele et al., Nature Scientific Reports, 2018).		5 60
➤ NEURON-TRACING-3D	Neuron tracing from 3D images. The images are from the DIADEM challenge (olfactory bulb projection fibers labeled with GFP) and were acquired by confocal microscopy (40x, NA = 1.3).		5 4
➤ NEURON-TRACING-TREES-3D	Neuron tracing in 3D images. The ground truth trees were generated by TREES Toolbox as SWC files and transformed into binary masks by Vaa3d . The masks were then convolved by a synthetic PSF (Born & Wolf) generated by ImageJ PSF Generator and some noise was added with ImageJ Random .		5 2
➤ NUCLEI-SEGMENTATION	Nuclei segmentation from 2D images. The images were generated by SIMCEP , a widefield fluorescence microscopy biological images simulator.		5 30

cytominE : visualize large images, remotely

- Explore **large** (multi-gigapixel) images at multiple resolutions, **remotely**
- OpenStreetMaps browsing style (zoom in/out, pyramid tile-based)

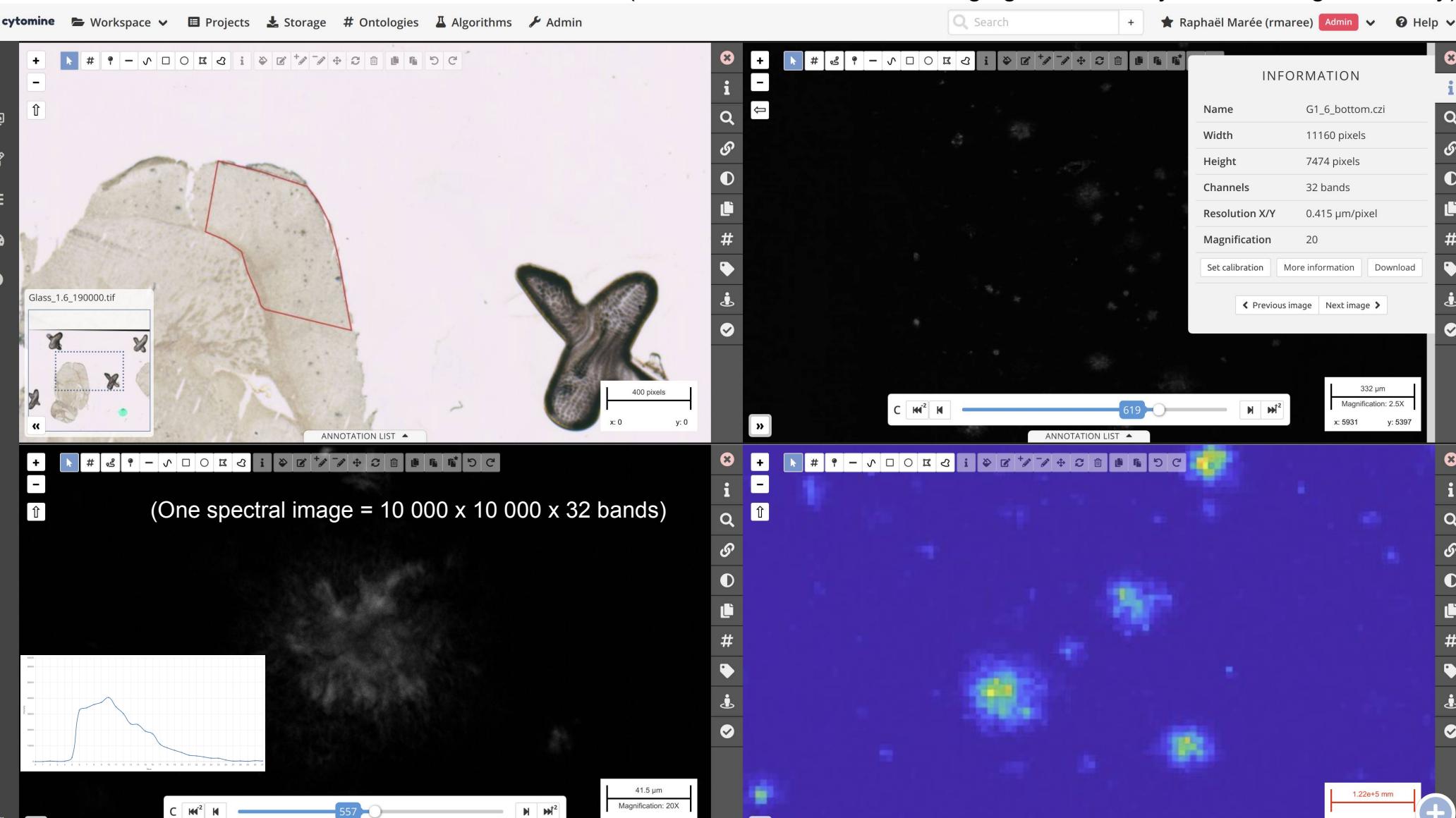


The screenshot shows the cytominE software interface. At the top, there is a navigation bar with links for 'cytominE', 'Workspace', 'Projects', 'Storage', 'Ontologies', 'Algorithms', and 'Admin'. On the right side, there is a user profile for 'Raphaël Marée (rmaree)' with an 'Admin' button. Below the navigation bar, there is a search bar and a help link. The main area of the interface displays a large tissue slice image. A text overlay on the left side of the image states: 'One tissue slice = 40000 x 30000 pixels (0.23µm/pixel)'. To the right of the main image, there is a zoomed-in view of a specific area, showing more detail of the tissue structure. The bottom left corner of the main image has a small thumbnail labeled 'PGP POUMON PB55 1 - 2012-08-07 11.39.07.jp2'. The bottom right corner of the zoomed-in image also has a similar thumbnail label.

... and multidimensional, multimodal images

- Explore **2D+c+z+t slices**
- Explore **spectral profiles**
- Side-by-side comparison of multiple views

(data: Centre for Cellular Imaging Core Facility, Gothenburg University)



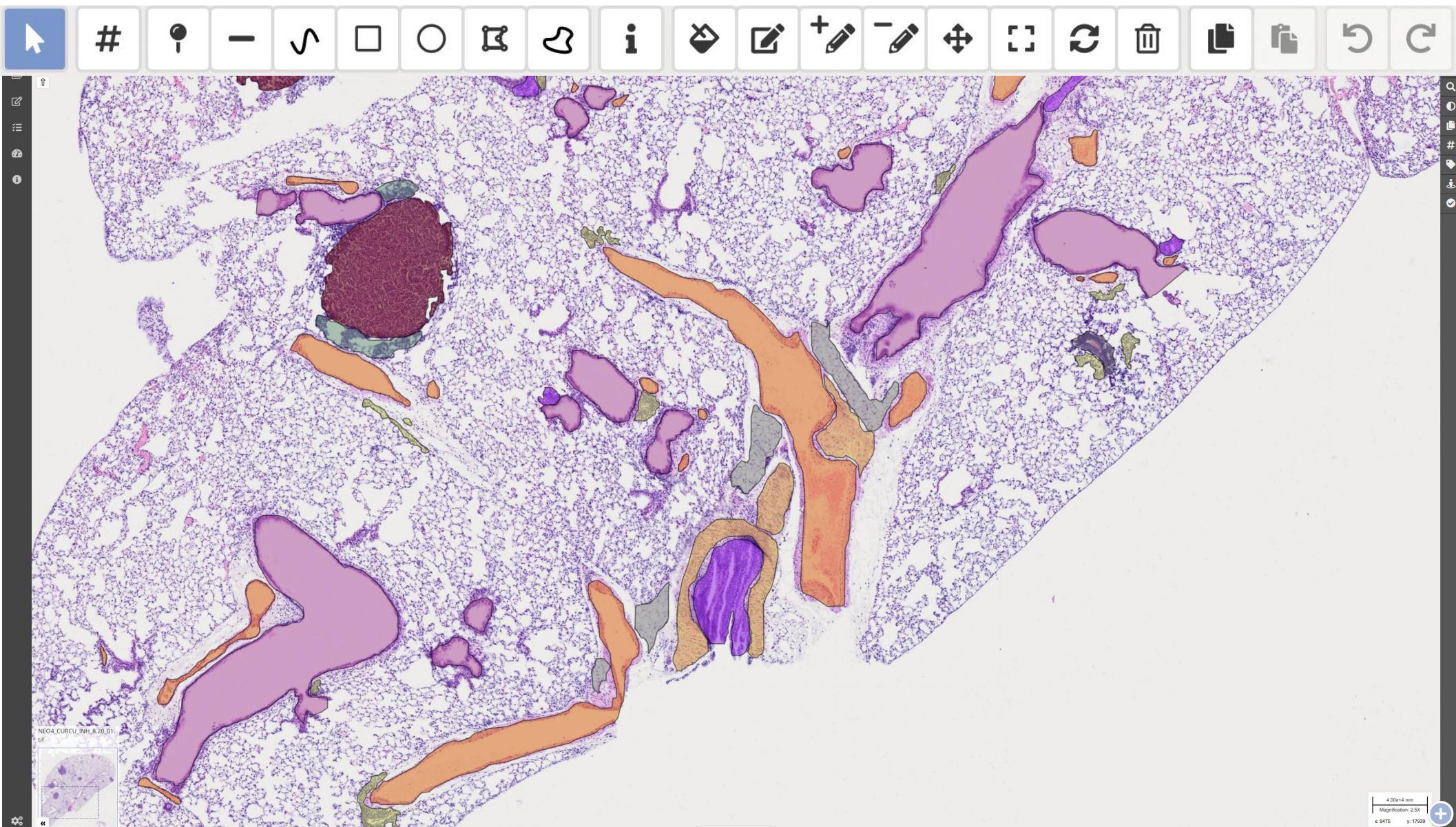
cytominE : annotate semantically, collaboratively

- Annotate images using various **drawing tools**, with **user-specific layers**
- Describe ROIs **semantically** with (user-defined) **ontology terms**
- Associate any **key-value properties, tags, file, or text description**



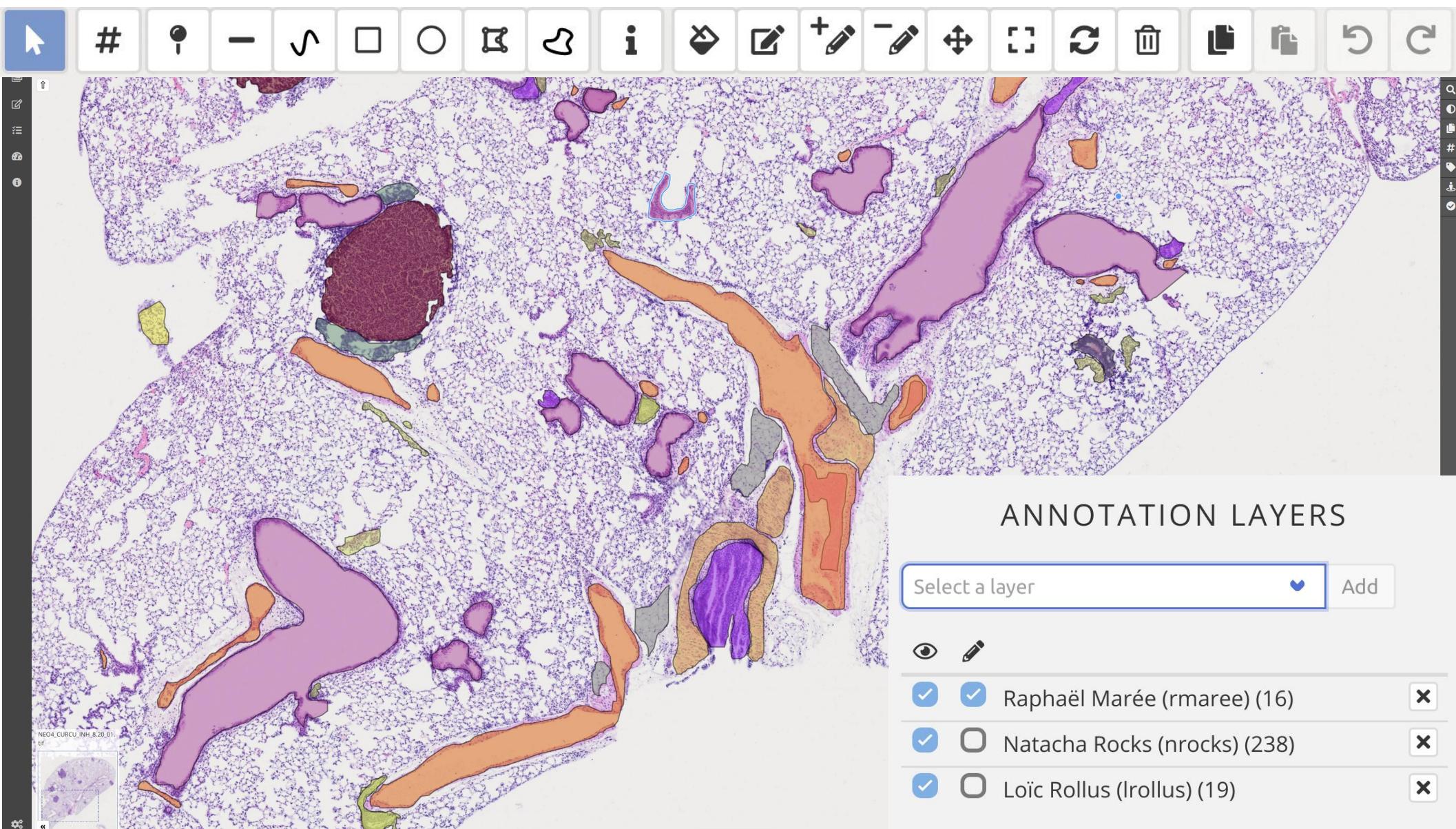
cytominE : annotate semantically, collaboratively

- Annotate images using various **drawing tools**, with **user-specific layers**
- Describe ROIs **semantically** with (user-defined) **ontology terms**
- Associate any **key-value properties, tags, file, or text description**



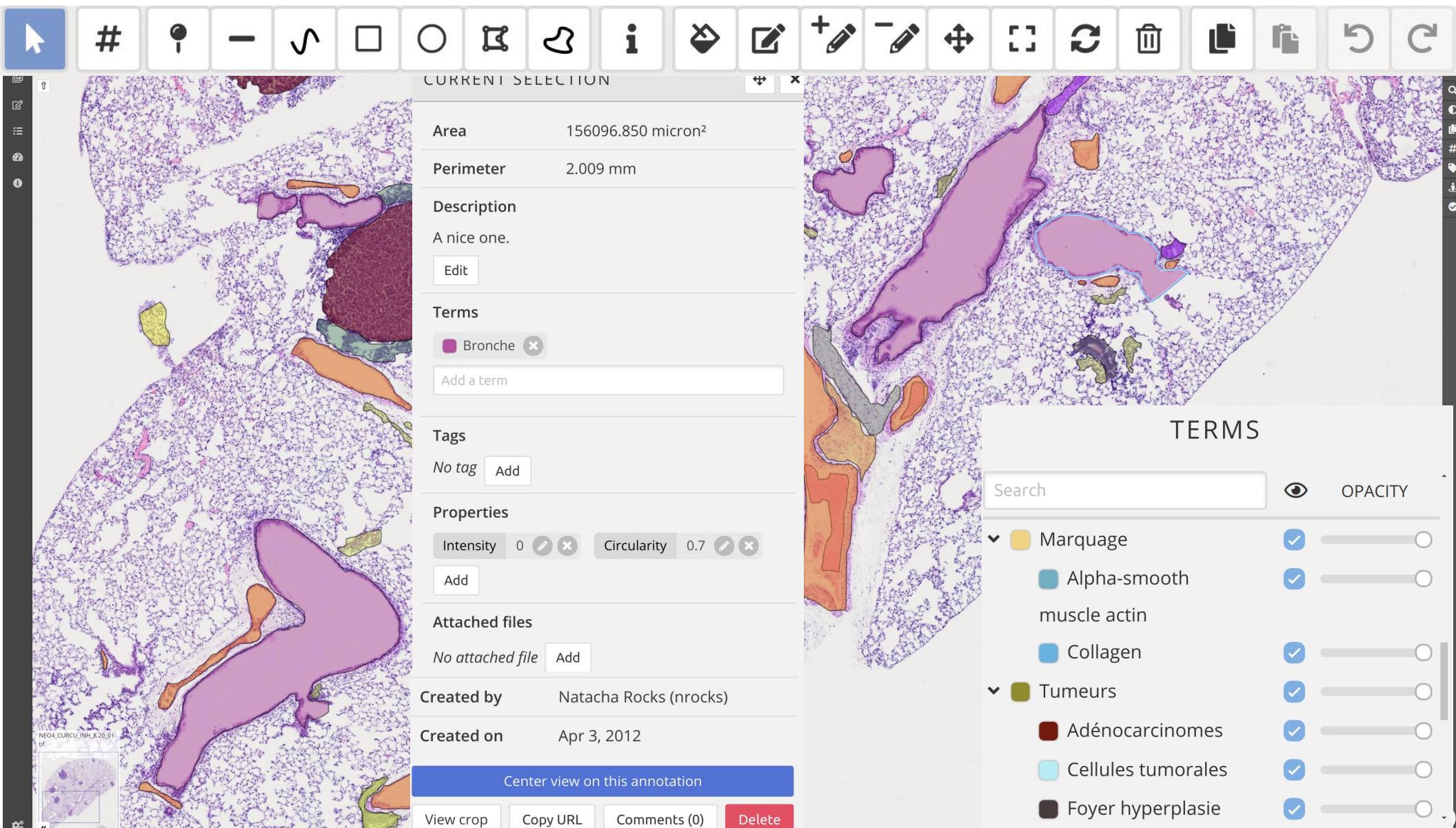
cytominE : annotate semantically, collaboratively

- Annotate images using various **drawing tools**, with **user-specific layers**
- Describe ROIs **semantically** with (user-defined) **ontology terms**
- Associate any **key-value properties, tags, file, or text description**



cytominE : annotate semantically, collaboratively

- Annotate images using various **drawing tools**, with **user-specific layers**
- Describe ROIs **semantically** with (user-defined) **ontology terms**
- Associate any **key-value properties, tags, file, or text description**



cytominE : annotate semantically, collaboratively

- Annotate images using various **drawing tools**, with **user-specific layers**
- Describe ROIs **semantically** with (user-defined) **ontology terms**
- Associate any **key-value properties, tags, file, or text description**

The screenshot shows the CytominE software interface for semantic annotation of histology images. At the top, there is a toolbar with various drawing tools and a status bar indicating the current selection and area (156096.850 micron²). A green notification box in the center says "Your comment was successfully send by email". Below it, the "Annotation comments" section displays two entries:

Raphaël Marée Nov 17, 2020 1:52 PM
This is another test comment.2.

Raphaël Marée Nov 17, 2020 1:52 PM
This is a test comment. How are you ?

At the bottom, there is an "Add comment" button, and a panel showing annotation details: Created by Natacha Rocks (nrocks), Created on Apr 3, 2012, and a "Center view on this annotation" button. The right side of the interface features a legend for tumor types and opacity controls.

cytominE : annotate semantically, collaboratively

Subject:Cytomine: Raphaël Marée (rmaree) commented an annotation

Date:Tue, 17 Nov 2020 12:52:57 +0000 (UTC)

From:cytominE.uliege@gmail.com

Reply-To:noreply@cytominE.org

To:info@cytominE.be



Hi,

Raphaël Marée (rmaree) shared an annotation with you and commented : "This is another test comment.2."

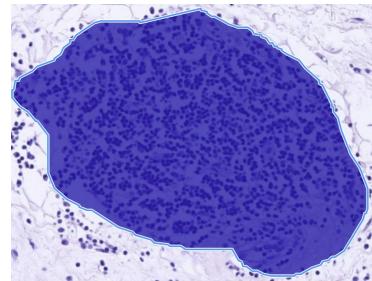


Navigate to <https://research.cytominE.be/#/project/51066666/image/51068020/annotation/51079563?action=comments> in order to reply.

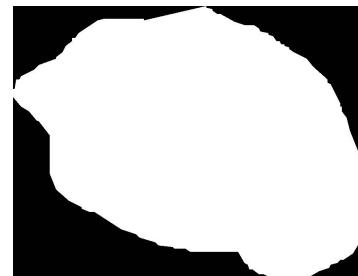
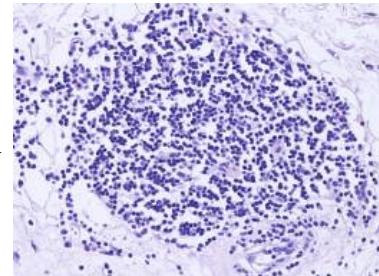
Navigate to <https://research.cytominE.be/#/project/51066666/image/51068020/annotation/51079563> in order to view the annotation within its context, or click on the thumbnail.

cytominE is highly interoperable

RESTful API to easily import/export all data through Internet



Web API



```
{"class":"be.cytomine.ontology.UserAnnotation","id":26675587,"created":15088778  
97417,"updated":null,"deleted":null,"slice":146698783,"image":24935980,"project"  
:"24796414","user":24768144,"location": "POLYGON ((40164 22674, 40322 22710, 40320  
22706, 40332 22702, 40336 22698, 40348 22690, 40392 22676, 40416  
22662, 40444 22662, 40446 22660, 40448 22658, 40450 22656, 40452 22654, 40454 22652,  
40456 22650, 40458 22648, 40508 22614, 40516 22614, 40516 22610, 40528  
22600, 40528 22600, 40540 22594, 40572 22578, 40568 22558, 40600 22520, 40644  
22518, 40652 22514, 40656 22466, 40656 22458, 40661 22498, 40660 22446, 40672  
22430, 40680 22398, 40704 22338, 40728 22302, 40744 22250, 40748 22198, 40744  
22186, 40744 22186, 40736 22170, 40688 22098, 40661 22074, 40656 22074, 40648  
22066, 40632 22062, 40628 22054, 40604 22046, 40576 22030, 40564 22030, 40560  
22026, 40500 22026, 40480 22036, 40464 22038, 40452 22038, 40436 22050, 40428  
22050, 40424 22062, 40416 22066, 40400 22094, 40280 22094, 40268 22102, 40240  
22102, 40236 22106, 40290 22119, 40192 22118, 40152 22130, 40144 22138, 40188  
22150, 40040 22194, 40028 22194, 40008 22202, 40008 22206, 39976 22222, 39944  
22250, 39928 22290, 39928 22384, 39908 22422, 39896 22422, 39876 22446, 39856  
22458, 39832 22482, 39836 22502, 39840 22502, 39844 22526, 39852 22526, 39856  
22534, 39884 22550, 39900 22562, 39944 22578, 39944 22582, 39960 22594, 39968  
22610, 39984 22622, 39984 22630, 39992 22630, 40000 22642, 40048 22674, 40064  
22678, 40164 22678, 40164 22674)"}, "geometryCompression": 3.0, "centroid":  
{"x": 40315.73937196692, "y": 22371.35279565198}, "area": 91131.05480176363, "areaUnit":  
"micron", "perimeter": 1.2623770183567887, "perimeterUnit": "mm", "term":  
"25635608], "nbComments": 0, "cropURL": "https://research.cytomine.be/api/  
userannotation/26675587/crop.png", "smallCropURL": "https://research.cytomine.be/  
api/userannotation/26675587/crop.png?maxSize=256", "url": "https://  
research.cytomine.be/api/userannotation/26675587/crop.png", "imageURL": "https://  
research.cytomine.be/#/project/24796414/image/24935980/annotation/  
26675587", "reviewed": false}
```

<https://research.cytomine.be/api/userannotation/26675587/crop.png?mask=true>

The screenshot shows the Icy software interface with three main panels:

- Left Panel:** Displays a microscopy image with a multi-colored overlay showing segmentation or detection results. It includes a zoom tool and coordinate/size information.
- Middle Panel:** Displays a microscopy image with a large white polygonal ROI. Numerous small colored crosses (red, blue, orange) are placed within the ROI, likely indicating specific features of interest. A magnification of 32.9X is shown. The status bar at the bottom indicates "Ready..." and provides coordinates (x=43547.29 px, y=35179.71 px).
- Right Panel:** An annotation viewer titled "Annotations - Cytomine". It lists a table of annotations with columns: Visible, Name, Terms, and Author. One row is highlighted in yellow: "106692475 [SubZoneTu... mboukerrou...". Below the table are buttons for "Associated terms", "Delete Annotation(s)", and "Set". At the bottom, there are summary statistics: ID 106599949, Dimension 24370.14 x 20497.70 µm, Magnification 40X, Annotations 0 (Algorithms), Annotations 142 (User), and Size 107136 x 90112 px.

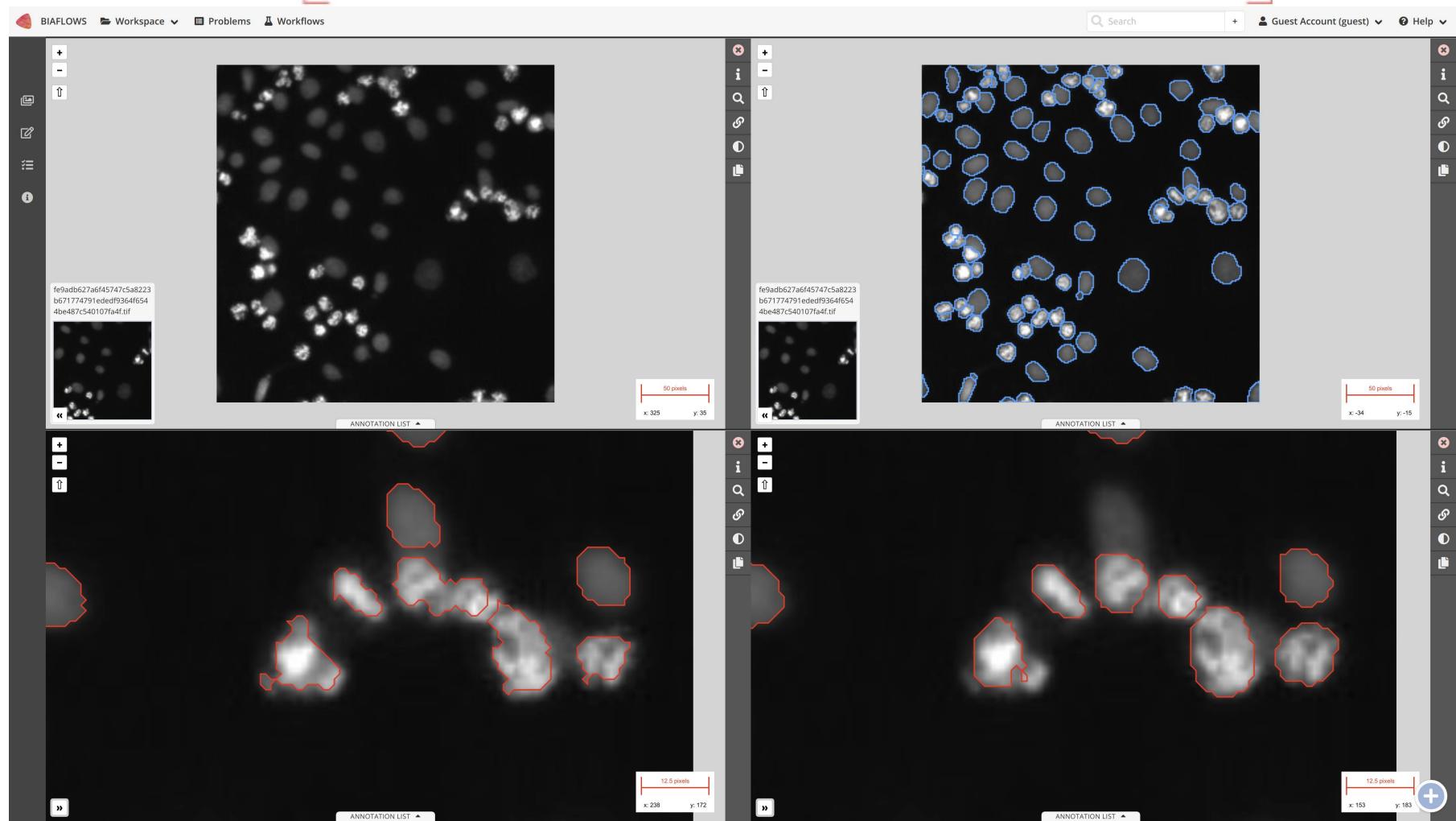


Institut Pasteur

cytominE is highly extensible

(Rubens et al., Cell Patterns, 2020)

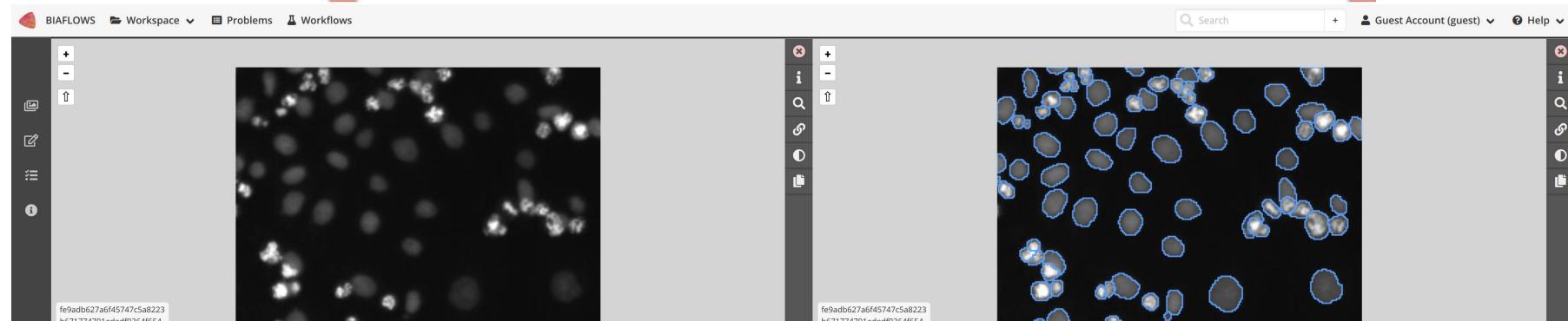
- **Integrate** algorithms **from any other tool** (ImageJ, Fiji, Icy, CellProfiler, ilastik, Vaa3D, Python, Keras/Tensorflow, OpenCV, ...) using **containers** (Docker/Singularity)
- **Reproducibility** : saving parameter values, versioned source code and libraries,...



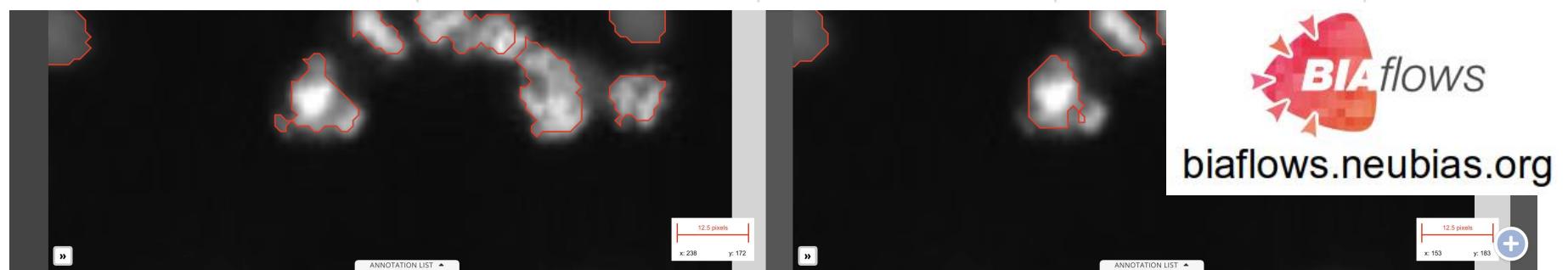
cytominE is highly extensible

(Rubens et al., Cell Patterns, 2020)

- **Integrate** algorithms **from any other tool** (ImageJ, Fiji, Icy, CellProfiler, ilastik, Vaa3D, Python, Keras/Tensorflow, OpenCV, ...) using **containers** (Docker/Singularity)
- **Reproducibility** : saving parameter values, versioned source code and libraries, ...



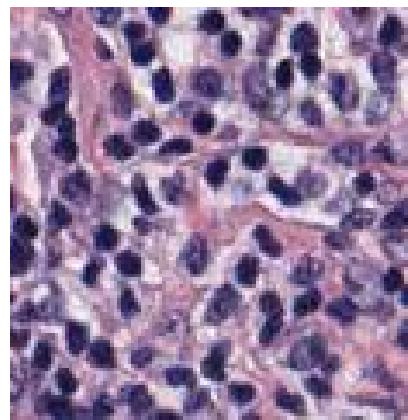
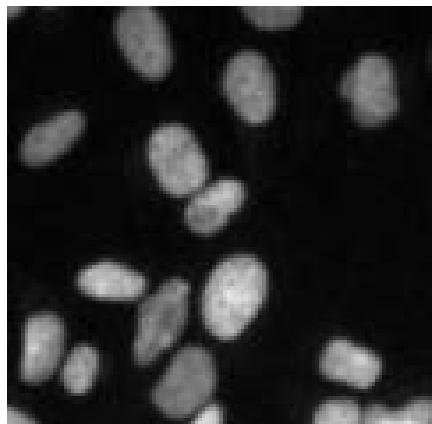
Workflow run	Mean Average Precision [Main metric]				Dice coefficient				Average Hausdorff distance				Fraction Overlap Pred			
	MIN	MAX	AVG	SD	MIN	MAX	AVG	SD	MIN	MAX	AVG	SD	MIN	MAX	AVG	SD
★ NucleiSegmentation-MaskRCNN (v1.5.0) #1 on Mar 28, 2020 7:49 PM	0.058	0.9	0.394	0.174	0.452	0.963	0.798	0.119	0.07	31.054	2.702	4.846	0.195	0.934	0.625	0.186
★ NucleiSegmentation-ilastik (v1.4.0) #2 on Mar 27, 2020 4:57 PM	0	0.82	0.208	0.195	0.215	0.944	0.725	0.164	0.056	34.331	3.843	6.527	0.046	0.902	0.502	0.181
★ NucleiSegmentation-UNet (v1.1.1) #7 on Mar 27, 2020 2:37 PM	0	0.85	0.282	0.177	0.075	0.948	0.754	0.193	0.053	128.964	8.485	20.924	0.044	0.934	0.542	0.177



An example of the potential benefits of an open science « sharing » approach

StarDist cell segmentation algorithm (Schmidt et al., 2018)

- Was trained on **public datasets** (images+ground-truths) : TCGA, DSB 2018 and Monuseg 2018 challenges
- It is **open-source** (<https://github.com/mpicbg-csbd/stardist>)
- It is a promising « generic » candidate (originally tested on fluorescent nuclei and H&E)



An example of the potential benefits of an open science « sharing » approach

StarDist integrated into cytominer

Launch new analysis

Algorithm CellDetect_Stardist_HE_ROI (v1.0.3) ▾

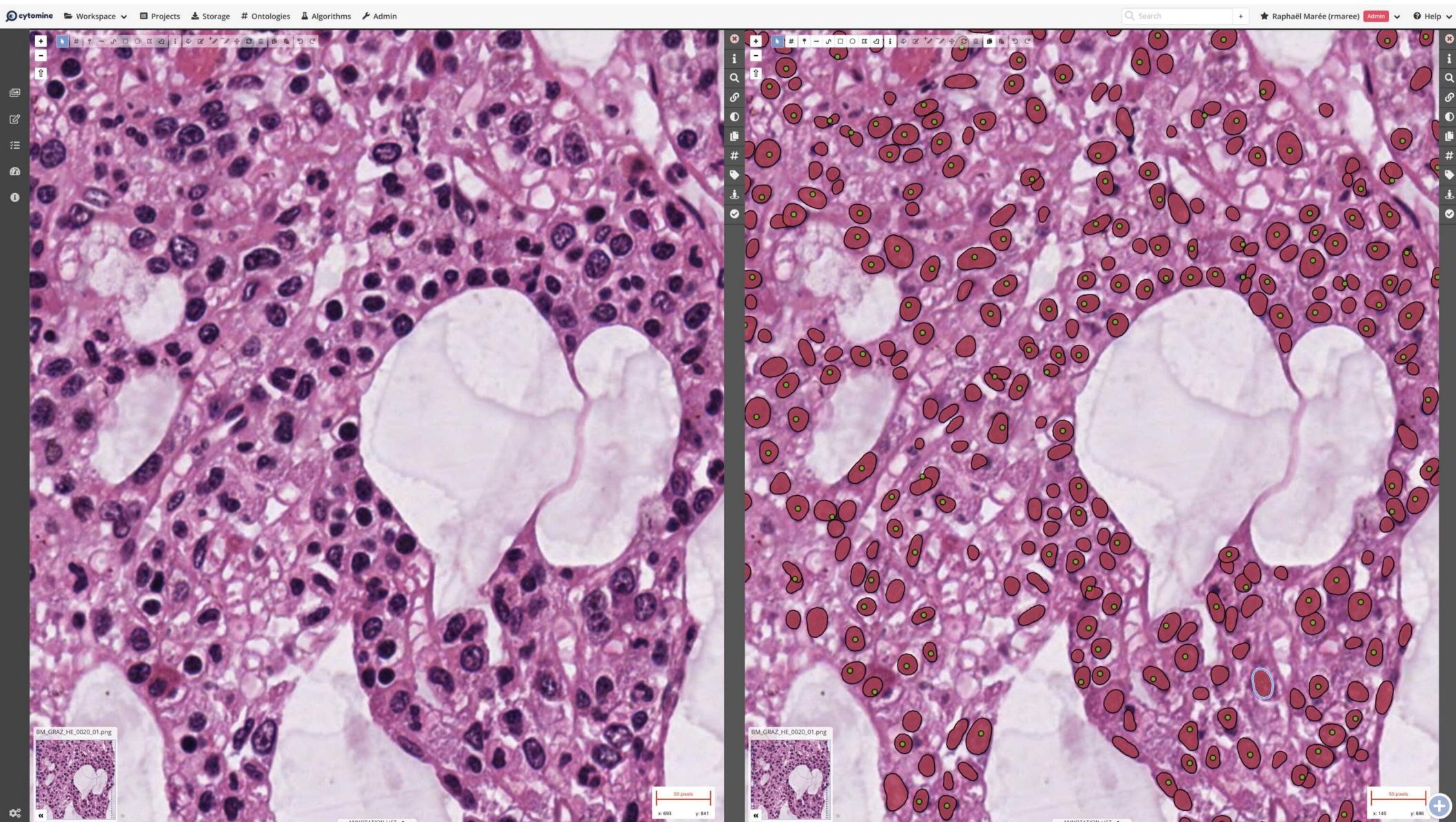
Name	Value
Cytomine Image IDs	CMU-1.svs
Cytomine ROI term ID	ROI
Cytomine Cell term ID	Nuclei_predict

Pre-filled parameters Hide

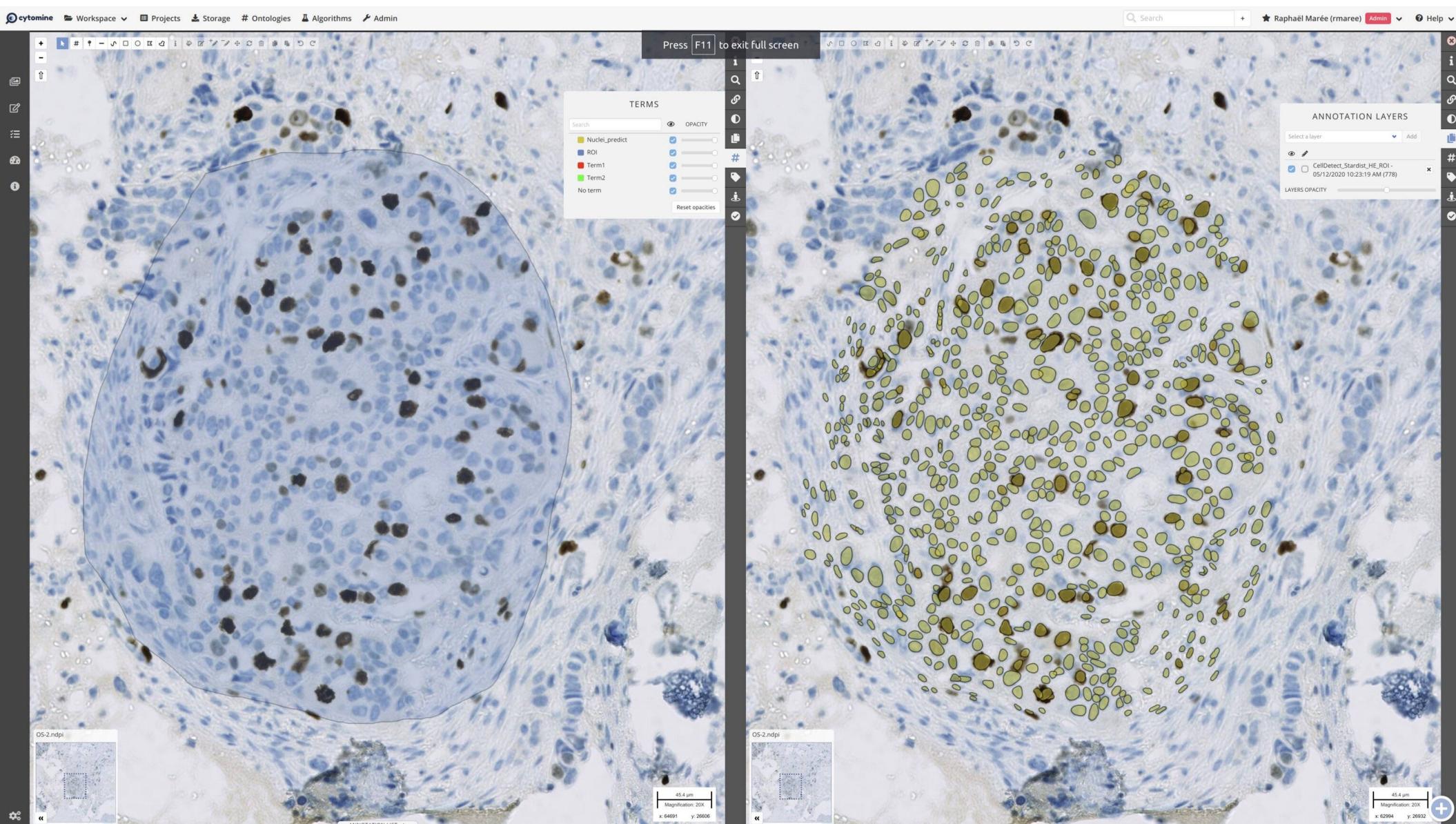
Stardist Probability Threshold	0.5
Stardist Non-Maximum Suppression Overlap threshold	0.5
Stardist Image Normalization Percentile Low	1
Stardist Image Normalization Percentile High	99.8

Cancel Launch new analysis

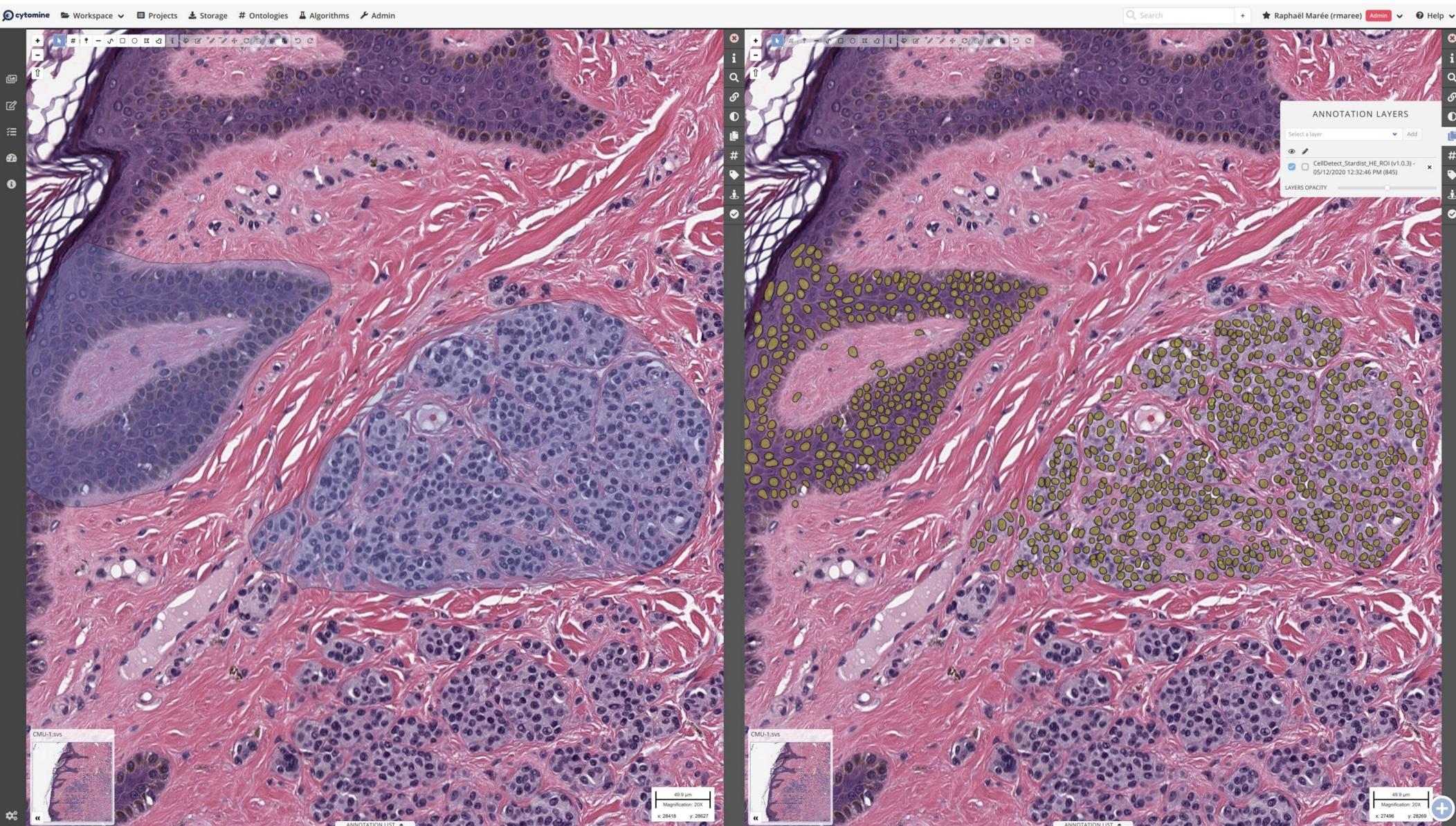
An example of the potential benefits of an open science « sharing » approach



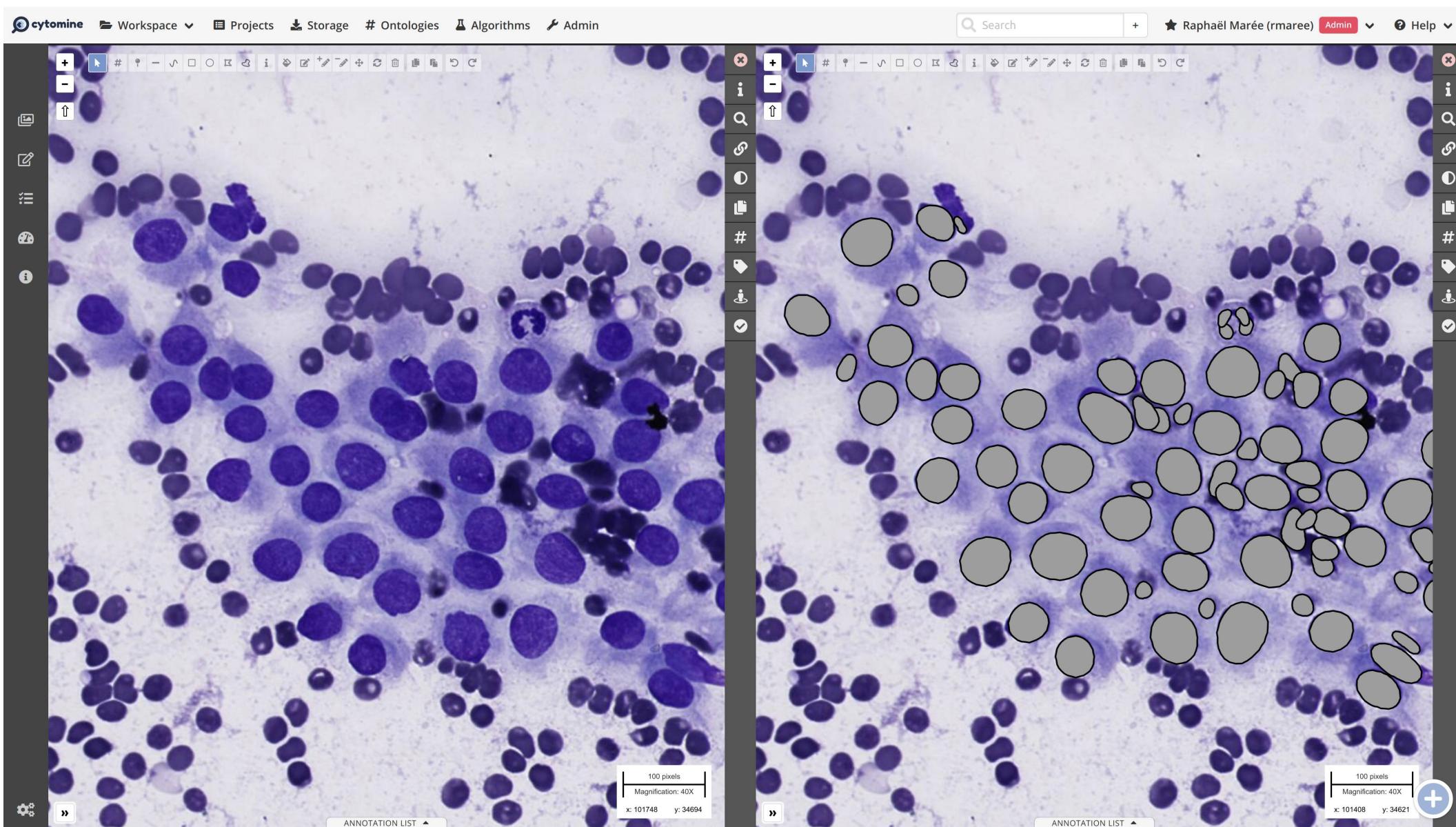
An example of the potential benefits of an open science « sharing » approach



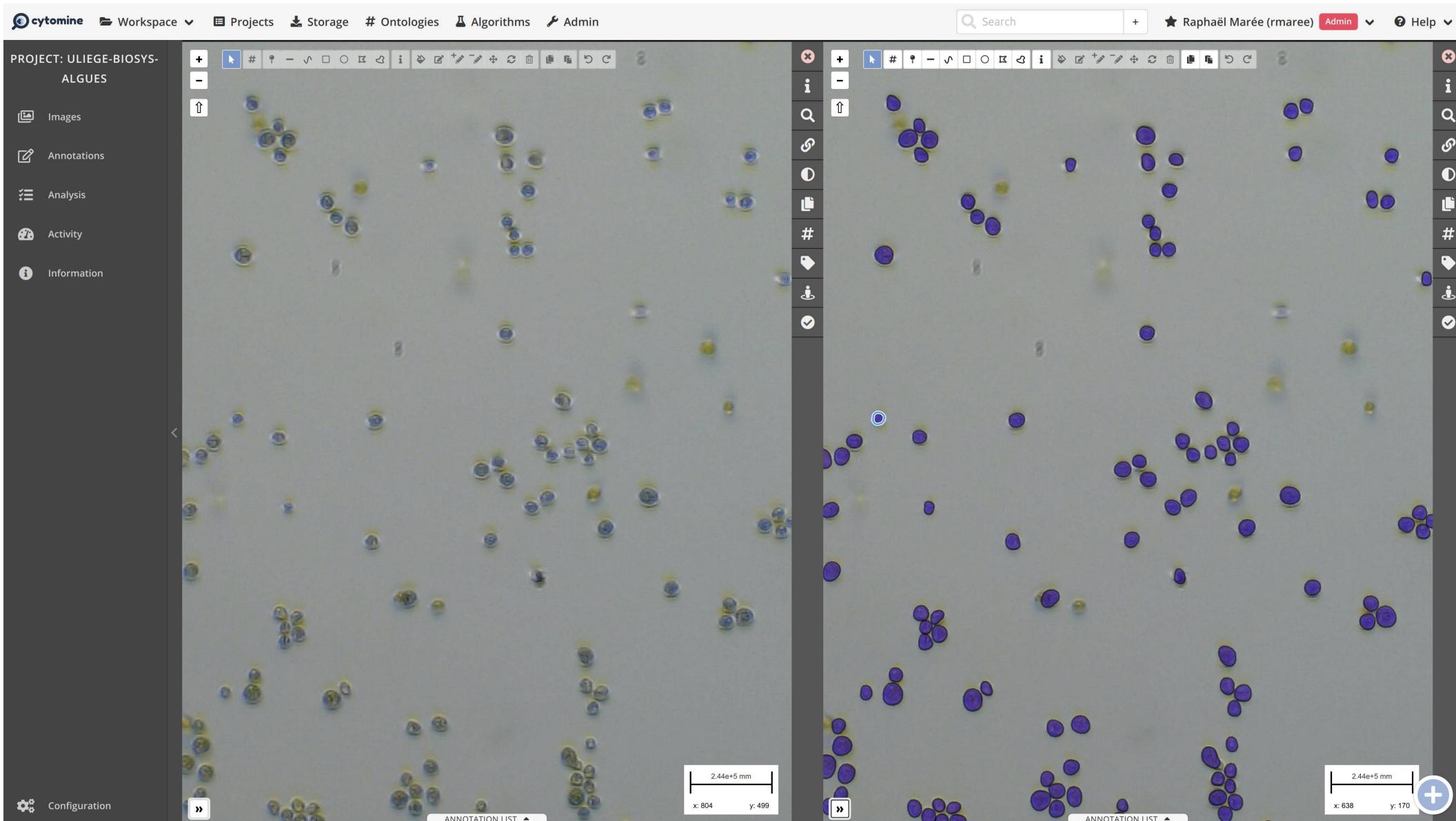
An example of the potential benefits of an open science « sharing » approach



An example of the potential benefits of an open science « sharing » approach



An example of the potential benefits of an open science « sharing » approach

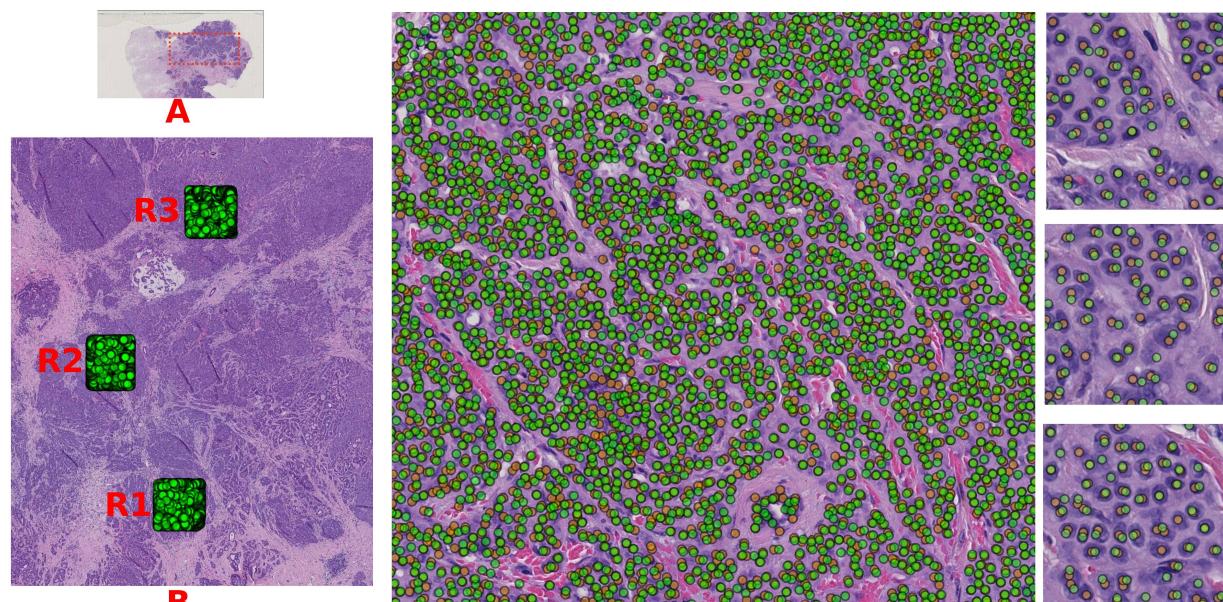


Qualitative vs Quantitative evaluation

Choosing an algorithm (e.g. for cell segmentation) only by **visual examination** of its predictions on a small subset of images is **unsafe** : no algorithm is perfect, bias, image variations and artefacts, ...

Ideally, **benchmarking** (quantitative evaluation) should drive image analysis method choice :

- Preselects methods that have good results on « similar », large datasets
- Evaluate quantitatively and tune methods on your own images + ground truths, proofread results



Reproducible, online, benchmarking



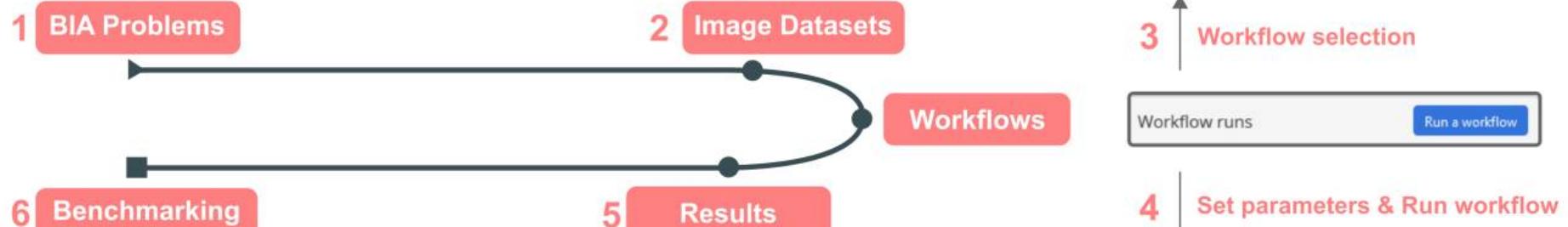
biaflows.neubias.org

NUCLEI-TRACKING-DIVISION
This project illustrates the 2D tracking of cell nuclei. The time-lapses are derived from Fluor-N2DH-SIM+ datasets from [Cell Tracking Challenge](#).

GLAND-SEGMENTATION-TRAIN
The images are crops of histopathology slides taken from the [2015 MICCAI challenge of gland segmentation](#) (GLaS 2015). The aim of the problem is to classify pixels as belonging to a gland or not. These images were used to train machine learning based workflows.

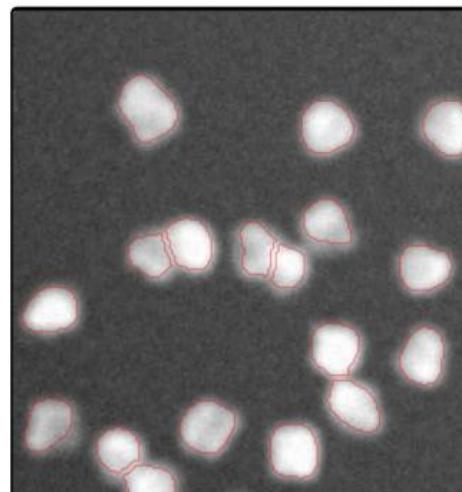
LANDMARKS-DROSO
Landmark detection in Drosophila wings, data from UPMC ([Vandaele et al., Nature Scientific Reports, 2018](#)).

VESSEL-TRACING-3D
This project illustrates the 3D tracing of blood vessels. The images were generated by [VascuSynth ITK](#), a biological image simulator, and some artificial noise was added.



Aggregated results Detailed results per image

Workflow run	Dice coefficient		
	MIN	MAX	AVG
★ NucleiSegmentation-ilastik (v1.0) #1 on Mar 25, 2019 1:19 PM	0.58	0.637	0.614
★ NucleiSegmentation-MaskRCNN (v1.3) #1 on Mar 25, 2019 9:16 AM	0.587	0.649	0.633
★ NucleiSegmentation-ImageJ (v1.10.1) #2 on Mar 19, 2019 10:37 AM	0.613	0.67	0.641
★ NucleiSegmentation-Python (v1.1) #6 on Mar 18, 2019 4:10 PM	0.554	0.613	0.586
★ NucleiSegmentation-CellProfiler (v1.4.1) #2 on Mar 11, 2019 9:22 AM	0.558	0.637	0.595



New workflow run

Workflow: Select options

- NucleiSegmentation-Python (v1.1)
- NucleiSegmentation-CellProfiler (v1.4.1)
- NucleiSegmentation-ImageJ (v1.10.1)
- NucleiSegmentation-MaskRCNN (v1.3)
- NucleiSegmentation-ilastik (v1.0)

Workflow runs Run a workflow

New workflow run

Workflow: NucleiSegmentation-ImageJ (v1.10.1)

Name	Value
Radius	5
Threshold	-0.5

Cancel Run a workflow