

Setting up mango in the lab for the research of neurodegenerative disease

Oct 2024

Mark Fiers

Lab

- PI: Bart de Strooper
- Solve Alzheimer's Disease
- Wet & Dry lab

The Major Risk Factors for Alzheimer's Disease: Age, Sex, and Genes Modulate the Microglia Response to A β Plaques

Carlo Sala Frigerio,^{1,2,3,*} Leen Wolfs,^{1,2} Nicola Fattorelli,^{1,2} Nicola Thrupp,^{1,2} Iryna Voytyuk,^{1,2} Inga Schmidt,^{1,2} Renzo Mancuso,^{1,2} Wei-Ting Chen,^{1,2} Maya E. Woodbury,⁴ Gyan Srivastava,⁴ Thomas Möller,⁴ Eloise Hudry,⁵ Sudeshna Das,⁵ Takaomi Saido,⁶ Eric Karran,⁴ Bradley Hyman,⁵ V. Hugh Perry,^{3,7} Mark Fiers,^{1,2} and Bart De Strooper^{1,2,3,8,*}

> [Cell](#). 2020 Aug 20;182(4):976-991.e19. doi: 10.1016/j.cell.2020.06.038. Epub 2020 Jul 22.




Spatial Transcriptomics and In Situ Sequencing to Study Alzheimer's Disease

Wei-Ting Chen¹, Ashley Lu¹, Katleen Craessaerts¹, Benjamin Pavie², Carlo Sala Frigerio³, Nikky Corthout², Xiaoyan Qian⁴, Jana Laláková⁴, Malte Kühnemund⁴, Iryna Voytyuk¹, Leen Wolfs¹, Renzo Mancuso¹, Evgenia Salta¹, Sriram Balusu¹, An Snellinx¹, Sebastian Munck², Aleksandra Jurek⁵, Jose Fernandez Navarro⁵, Takaomi C Saido⁶, Inge Huitinga⁷, Joakim Lundeberg⁵, Mark Fiers⁸, Bart De Strooper⁹

Me

- Bioinformatics
- Teach bioinformatics
- PhD & Master students

Novel Alzheimer risk genes determine the microglia response to amyloid- β but not to TAU pathology

Annerieke Sierksma¹⁰, Ashley Lu, Renzo Mancuso¹⁰, Nicola Fattorelli¹⁰, Nicola Thrupp, Evgenia Salta, Jesus Zoco¹⁰, David Blum¹⁰, Luc Buée, Bart De Strooper¹⁰  , and Mark Fiers¹⁰  | [AUTHOR INFORMATION](#)

EMBO Mol Med(2020)12: e10606 | <https://doi.org/10.15252/emmm.201910606>

Stem-cell-derived human microglia transplanted in mouse brain to study human disease

[Renzo Mancuso](#) , [Johanna Van Den Daele](#), [Nicola Fattorelli](#), [Leen Wolfs](#), [Sriram Balusu](#), [Oliver Burton](#), [Adrian Liston](#), [Annerieke Sierksma](#), [Yannick Fourné](#), [Suresh Poovathingal](#), [Amaia Arranz-Mendiguren](#), [Carlo Sala Frigerio](#), [Christel Claes](#), [Lutgarde Serneels](#), [Tom Theys](#), [V. Hugh Perry](#), [Catherine Verfaillie](#), [Mark Fiers](#) & [Bart De Strooper](#) 

[Nature Neuroscience](#) **22**, 2111–2116 (2019) | [Cite this article](#)

Wet & Dry lab

- (relatively) data heavy techniques
 - Transcriptomics (spatial, single cell, bulk)
 - Genetics & genomics
 - Image analysis
 - 3rd party data analysis
- Very **heterogeneous** – all projects are different
 - Different methods, formats, tools, biology, bioinformaticians, collaborations, biologists
 - Time pressure

Premise

We're not experts in this. We do need to track our data (reproducible science). We don't have time

- Probably reinventing a few wheels
- Must integrate with our ELN (automatically)
- Needs to fit with existing workflows
- Current solution is hacky.
 - Python/click based command line workflow
- Open to suggestions

Old (non-)solution

- Jupyter notebooks
- At discretion of the data analyst
 - Weak directory project structure, across multiple servers
 - Sub optimal (or horrible) lab journal software
 - Note taking in obsidian (again not structure enforced)
- Data is all over the place
 - Backup share organized by Raf
 - We did start to use Raf's ingress

Elabjournal

- Web based
- More user friendly
- Project
 - Study
 - Experiment



eLabJournal

Journal | Inventory | Protocols | Supplies | Configuration | File Storage | Marketplace

Dashboard | Experiment Browser | Timeline | Projects | Studies | Experiment list

Experiment Browser

Search by project, study or experiment...

- Bart De Strooper lab

- Ck1delta Project
- + Demo
- + Sex Chromosomes
- + Test

- Bioinformatics

- 101010 project
 - Visium Spatial Transcriptomics 3
 - Single cell transcriptomics 1
 - Xenium Spatial Transcriptomi... 1
- + Lab data archive

Experiment Name	Status
Differential Expression Models	Configuring
Villages 3M	Configuring
City Villages at 6M	Configuring
Single grafts	Configuring

elabjournal



- Web based
- More user friendly
- And, an API (could be better documented)

Experiment

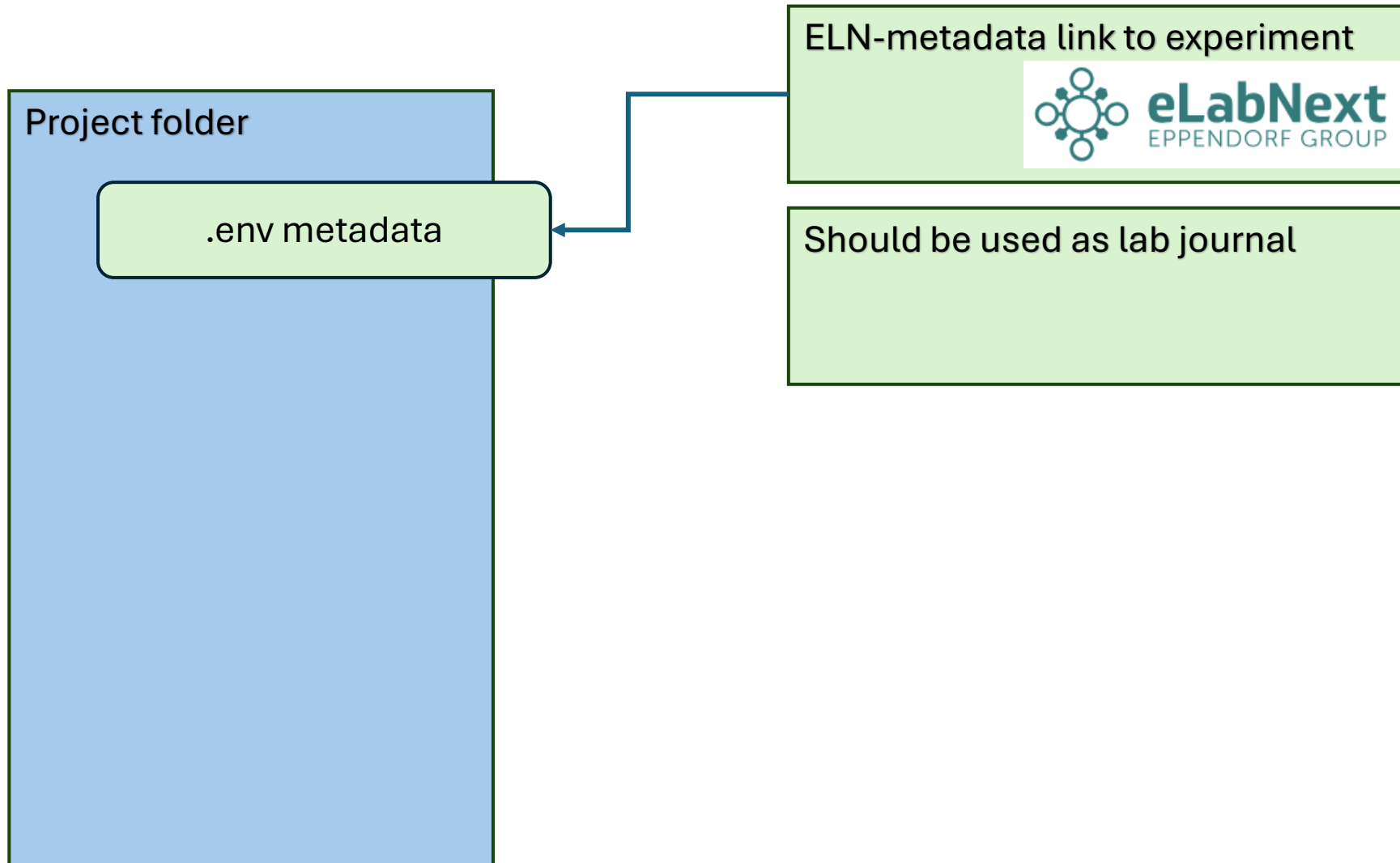
Show/Hide | List Operations | Expand Operations

GET	/api/v1/experiments/{experimentID}/signatureWorkflow	Get the signature workflow for an experiment
GET	/api/v1/experiments	Get experiments
POST	/api/v1/experiments	Create a new experiment
GET	/api/v1/experiments/{experimentID}	Get an experiment by id
GET	/api/v1/experiments/{experimentID}/logs	Get an experiment's change logs.

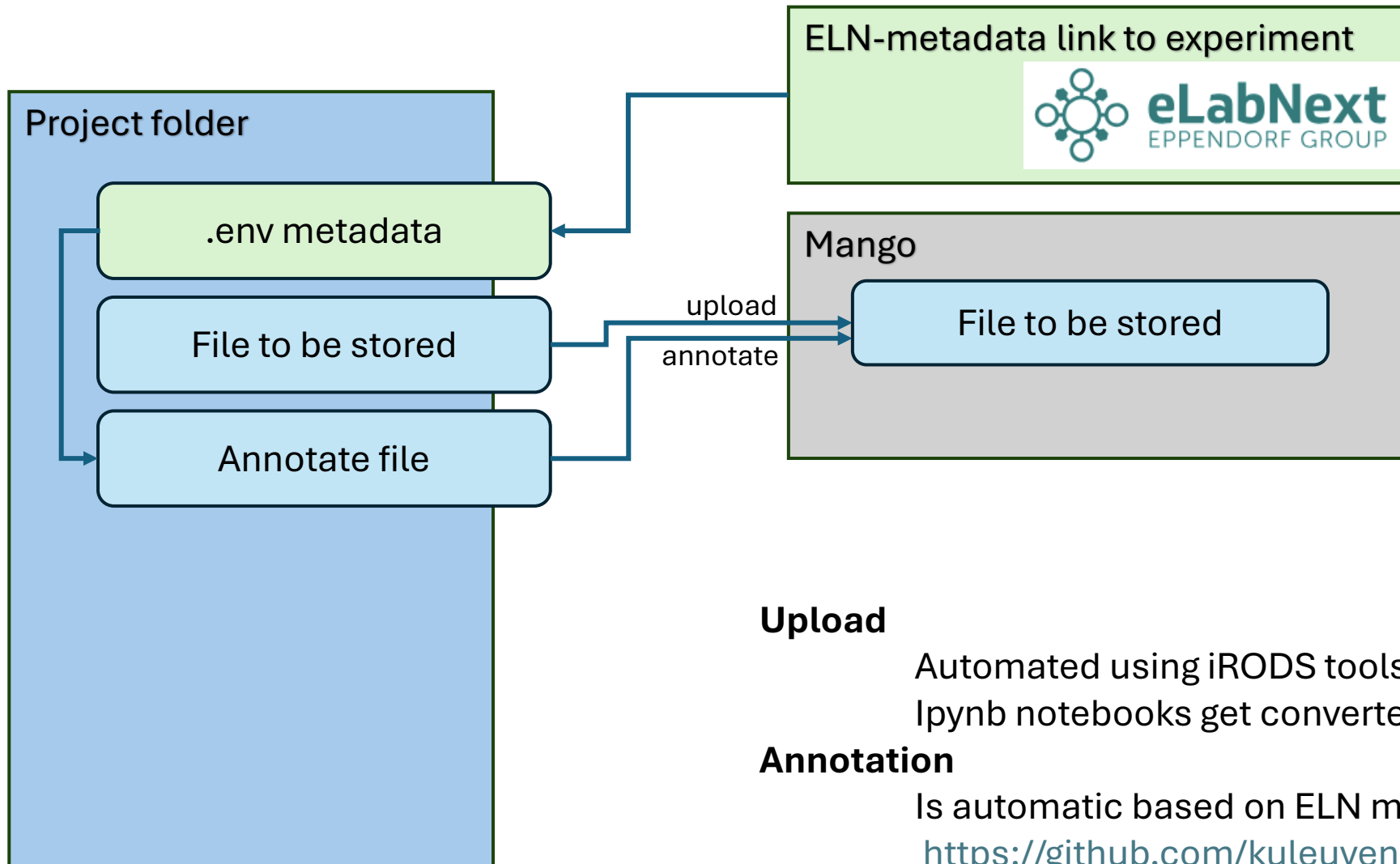
Current (proposed) workflow

- ELN to track projects
 - Use ELN to track files stored in Mango
 - Mango metadata contains ELN IDs
 - Link work folders to ELN experiments
- Friendly to bioinformaticians
 - Don't disrupt workflow
 - Command line
 - Use ipynb or Rmd notebooks
- Partly in place

Current (proposed) workflow



Current (proposed) workflow



Upload

Automated using iRODS tools or Gocommands.
Ipython notebooks get converted to pdf & uploaded to ELN

Annotation

Is automatic based on ELN metadata using
<https://github.com/kuleuven/mango-mdschema>



version 7.0.0

published

View

New (draft) version

Copy to new schema

Download JSON

Archive

Study name*

Input type: text

Project name*

Should correspond to elabjournal project name

Project id*

Input type: integer

ELN ID

Server*

Input type: text

Study id*

Input type: integer

ELN ID

Experiment name*

Short experiment name

Experiment id*

Should be a number like - 1289712

ELN ID

Collaborator*

Input type: text

Path*

Input type: text

Upload date*

Input type: date

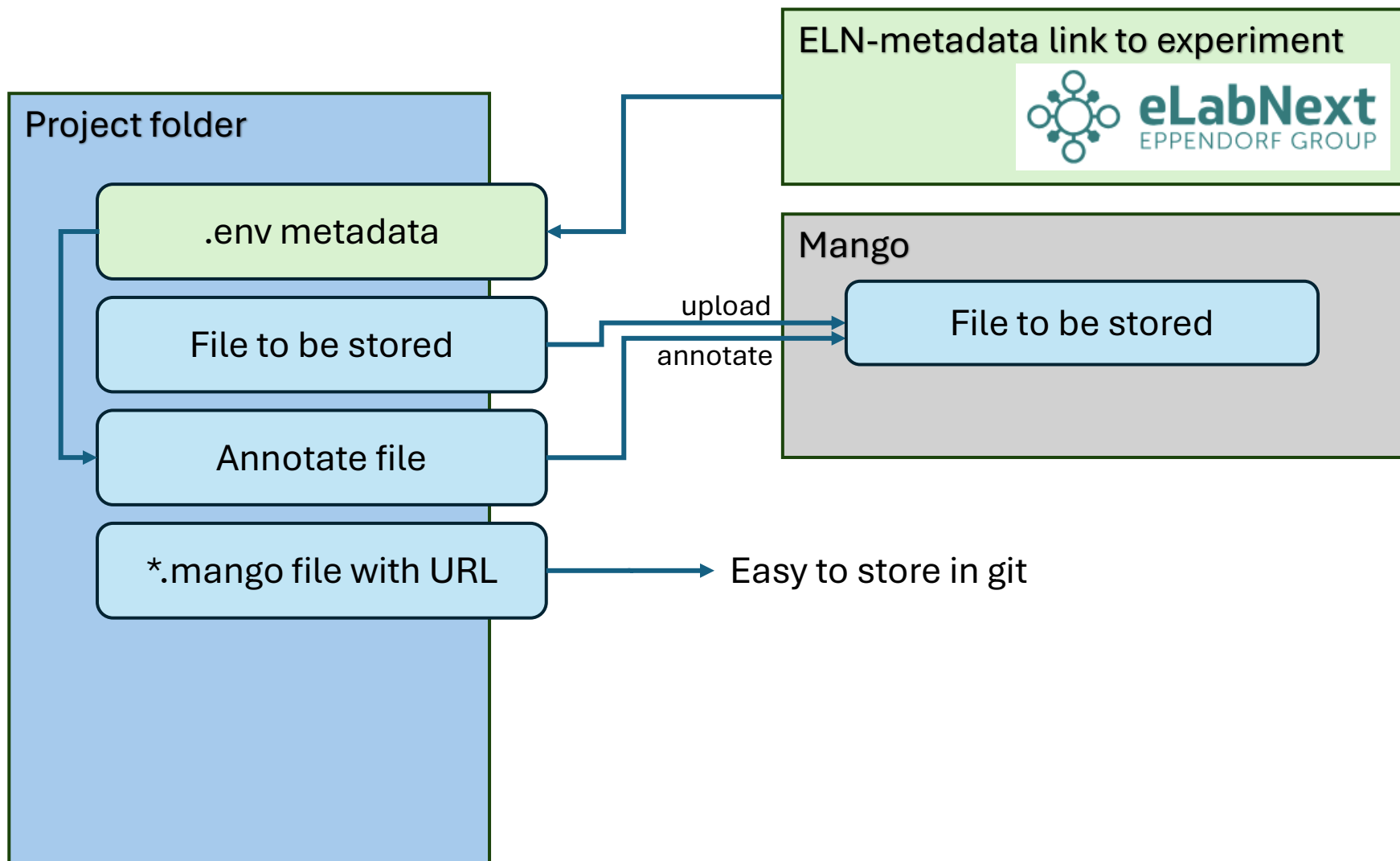
Description

Input type: textarea

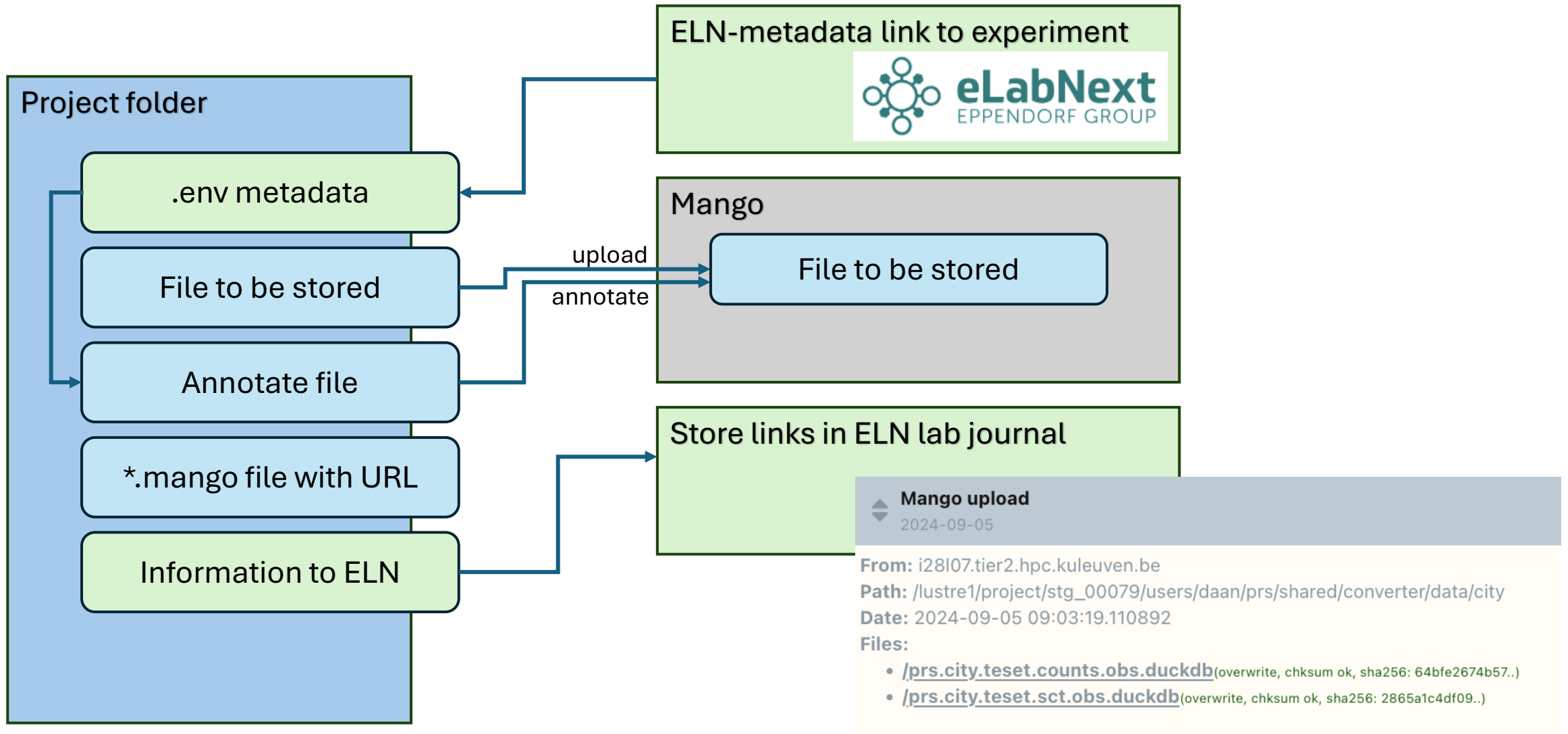
Sha256 Checksum*

Input type: text

Current (proposed) workflow



Current (proposed) workflow



Some observations

- Using metadata
 - I really like versioned metadata schema's & the ability to automatically generate them
 - Need to learn querying
- I like the globus backend
 - Easy for moving files around

Some observations

- Interface with iRODs / Mango?
 - Standard irods tools?
 - How well maintained is this?
 - Did not get it to work on my Mac/M1
 - Python - <https://github.com/irods/python-irodsclient>
 - Appears slow for many large files?
 - Does not seem to work on a mac/M1?
 - Better? Gocommands (in Go, works on a Mac)
 - <https://github.com/cyverse/gocommands>
 - Aim to migrate to Gocommands

Conclusion – I like Mango

- Workflow is hacky, but works
- API's allow adapting the workflow to us (instead of vice-versa)
- Fine balance between a workable system and a straightjacket

Wishlist

- Cold storage!
 - Lots of data of old projects that need storage
- Deeper integration with workflows?
 - E.g. Snakemake
 - Git-lfs backend?
- Get our team to use this consistently
 - Lab journal is a bigger problem

Teaching

- With Ingrid & Mariana
- Introduction to mango
 - Store & retrieve data
 - Annotate & search metadata

Future

- Deeper embedding of Mango
- Integrate in workflows?

Management of Large-Scale Omics Data [I0U19a]

*Students will be able to design and use **systems to store and serve large biological datasets**, and to obtain disparate data from online resources. They will have an up-to-date view of technologies that are used for storing, exchanging, processing and querying of big datasets.*

