

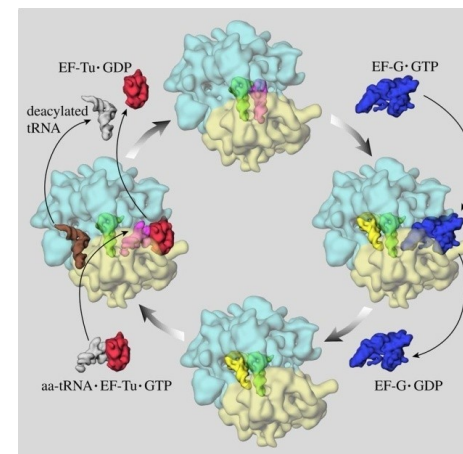
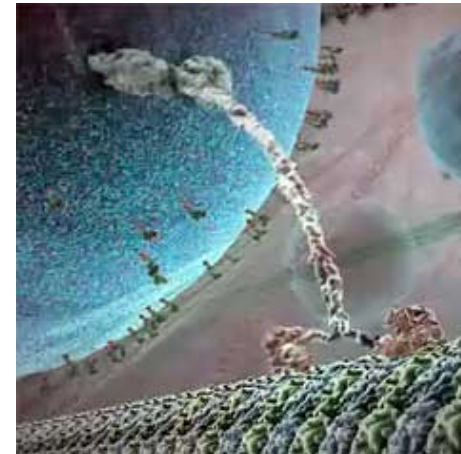
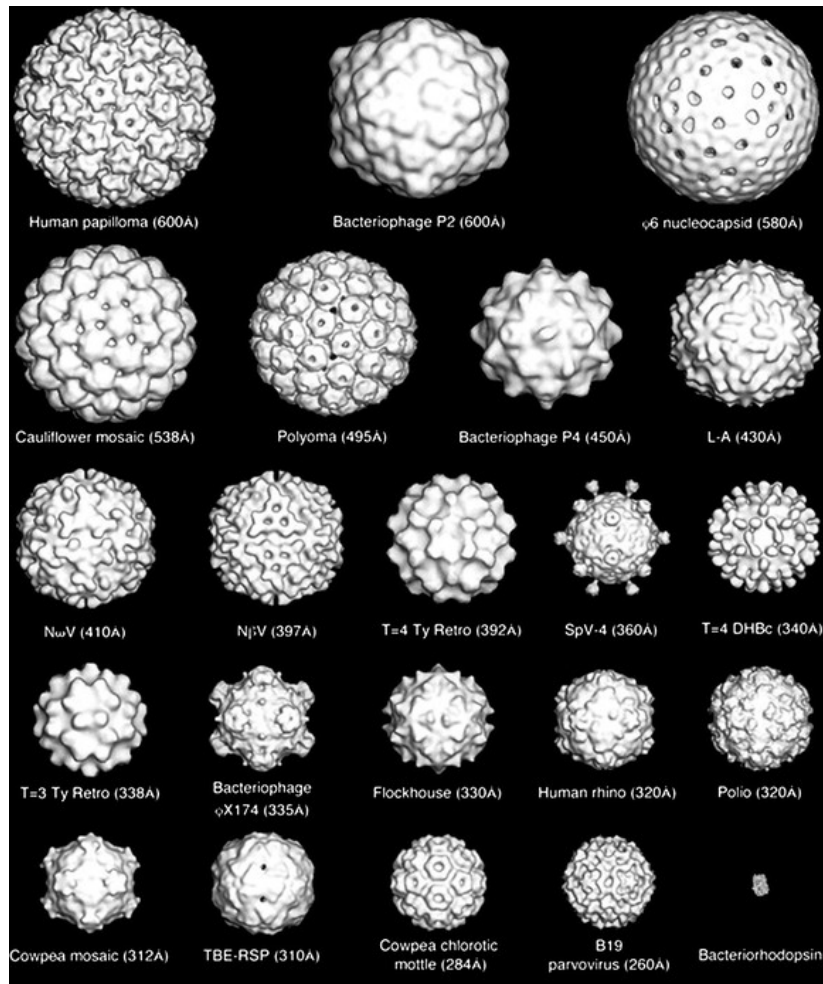


# Facing Computing and Data Challenges in Electron Microscopy: The **SCIPION** Software Framework

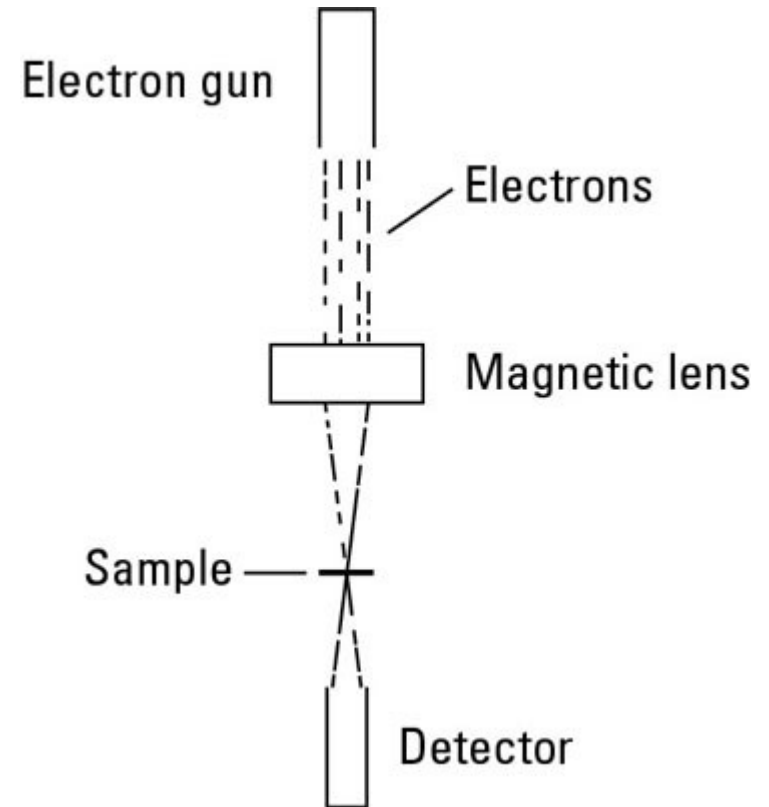
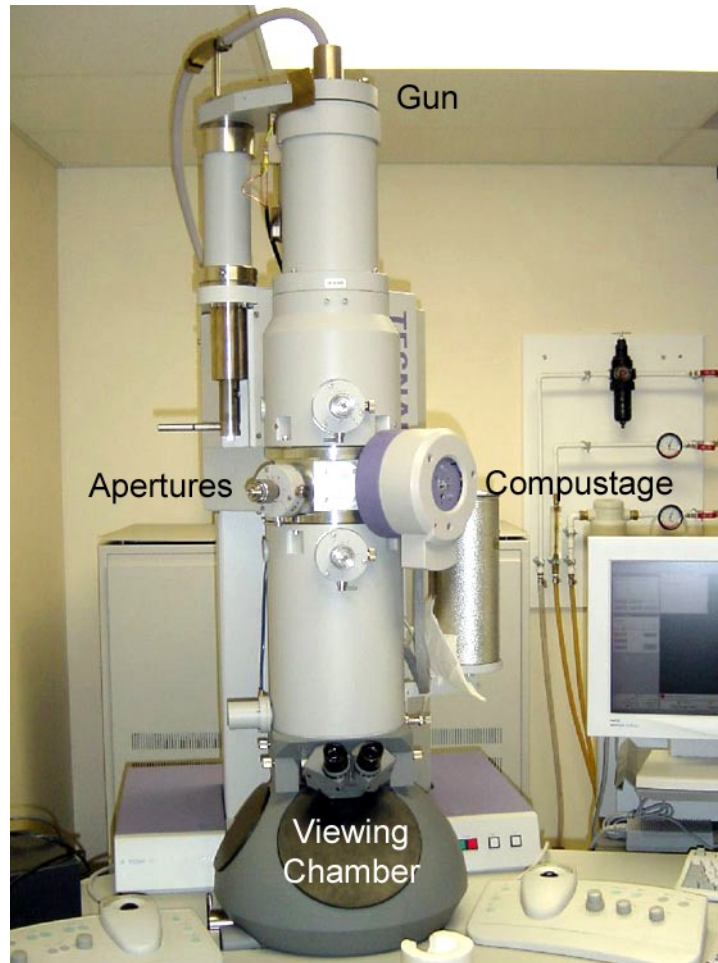
Biocomputing Unit,  
Instruct Image Processing Center, CNB-CSIC  
J.M. de la Rosa Trevín



# Life is based on Molecular Machines



# Our microscope use electrons to form images

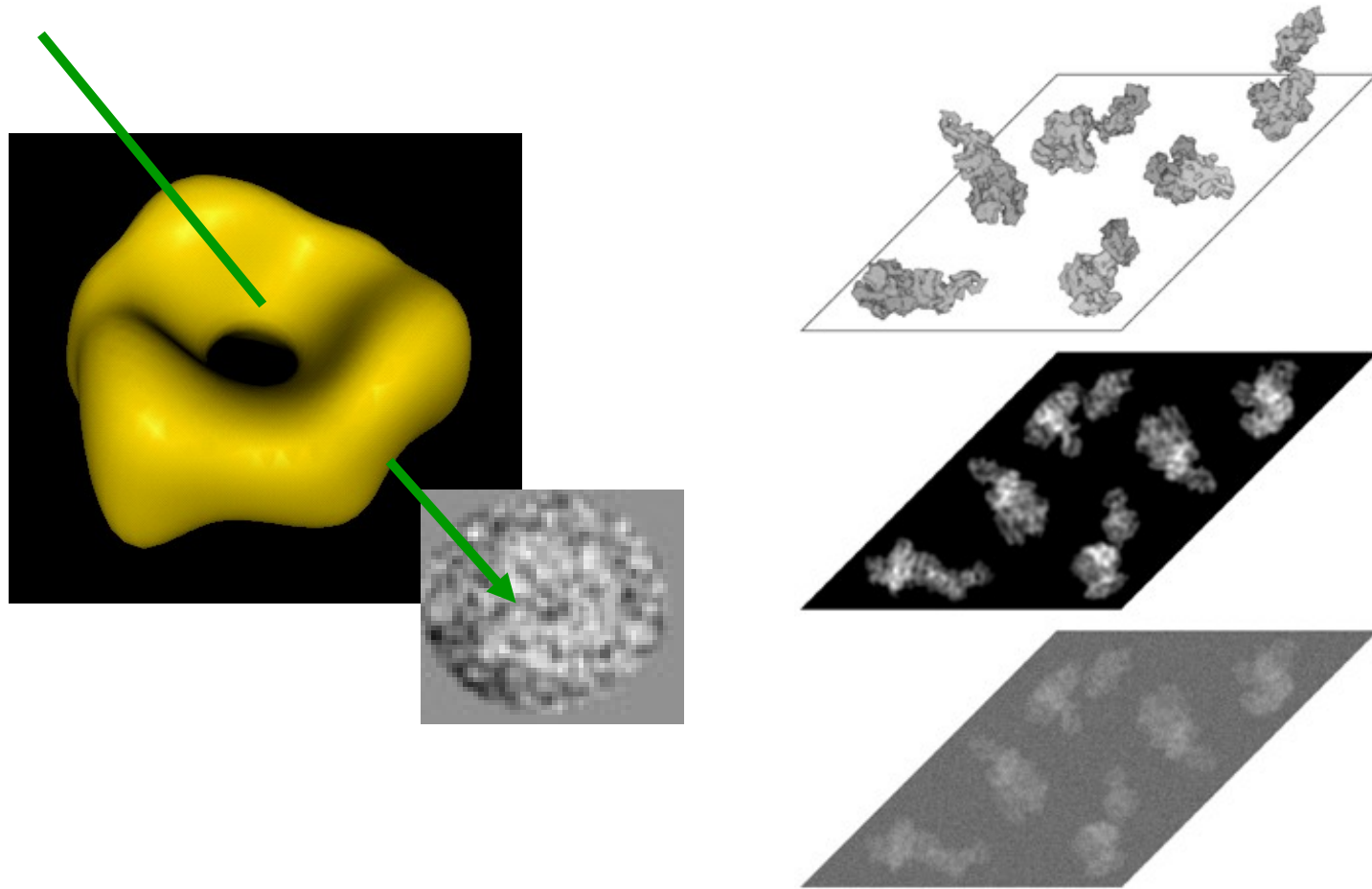


# Microscopy can be fancy (and expensive!)

---



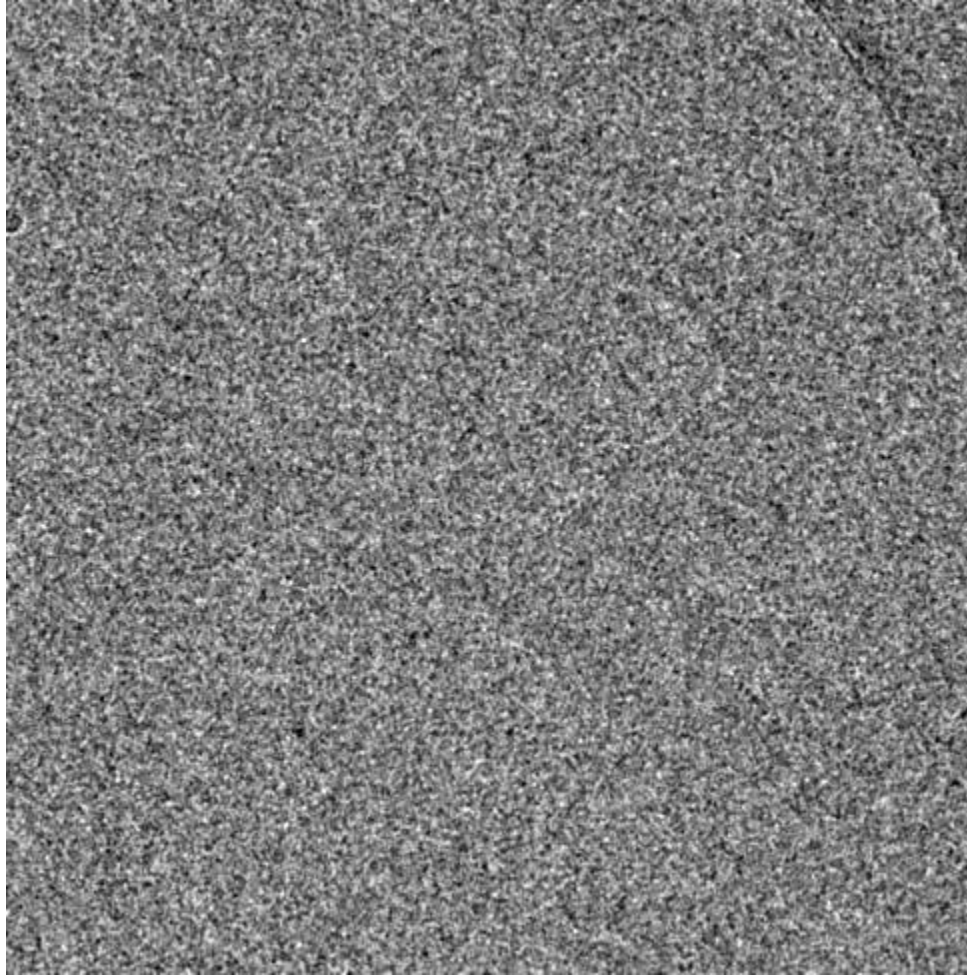
# We use a low electron dose





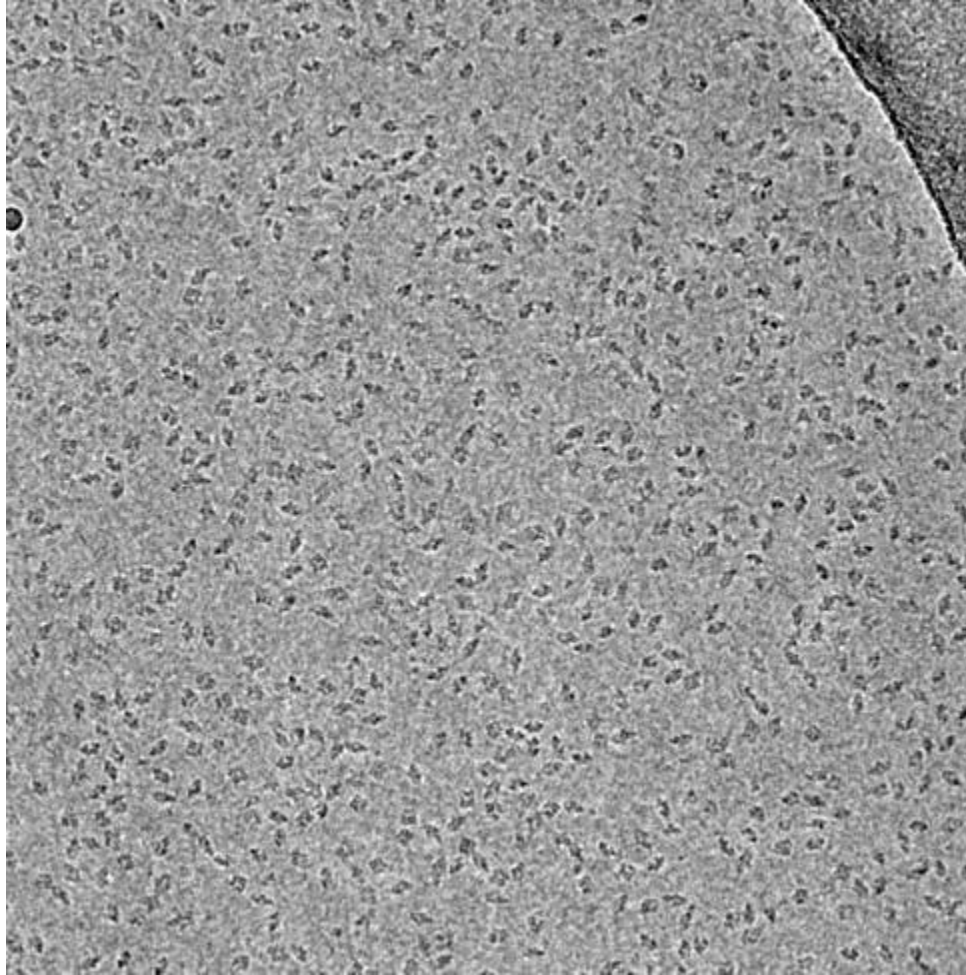
# EM-images biological are extremely noisy

---



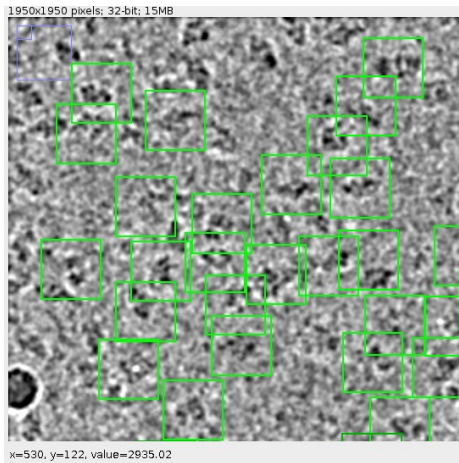
# We need to 'average' to extract some signal

---

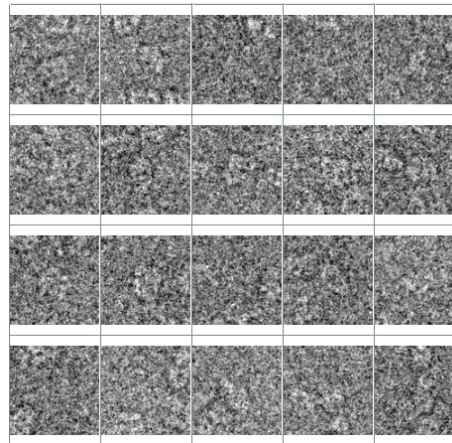


# A long path from 2D images to 3D volumes

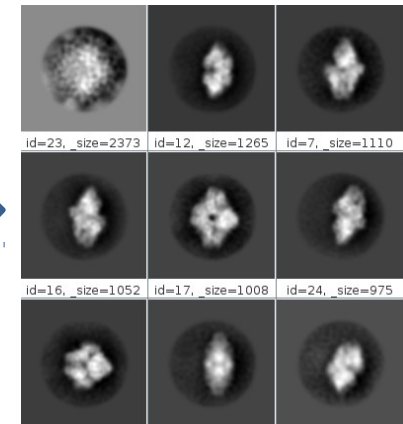
Particles selection



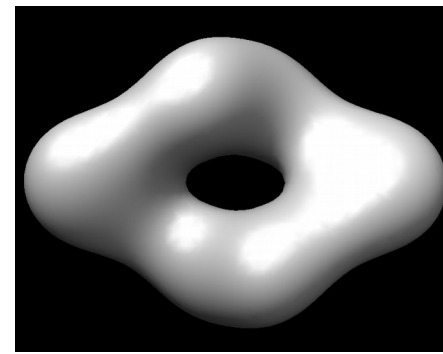
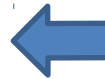
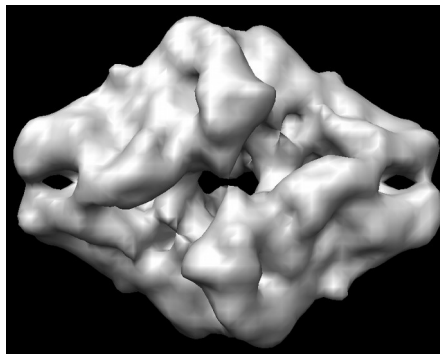
Extraction



2D classification



Final 3D map



Initial volume



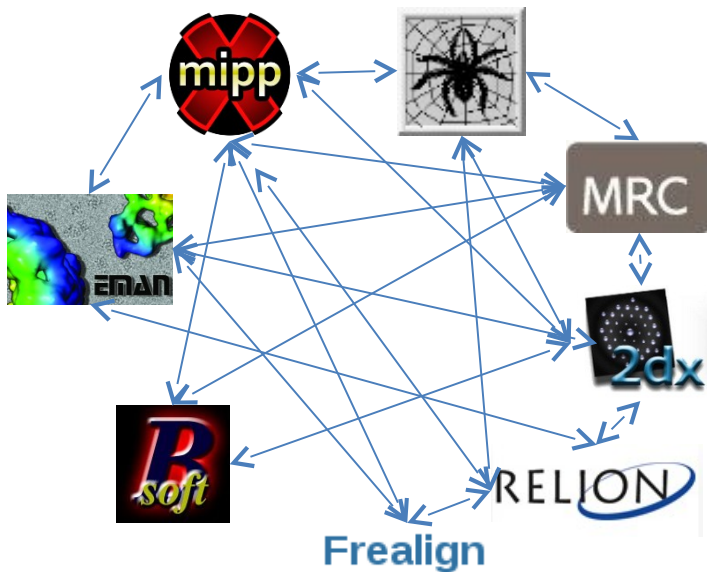


# Software Challenges: Interoperability, Reproducibility and Automation



# The EM field needs software integration

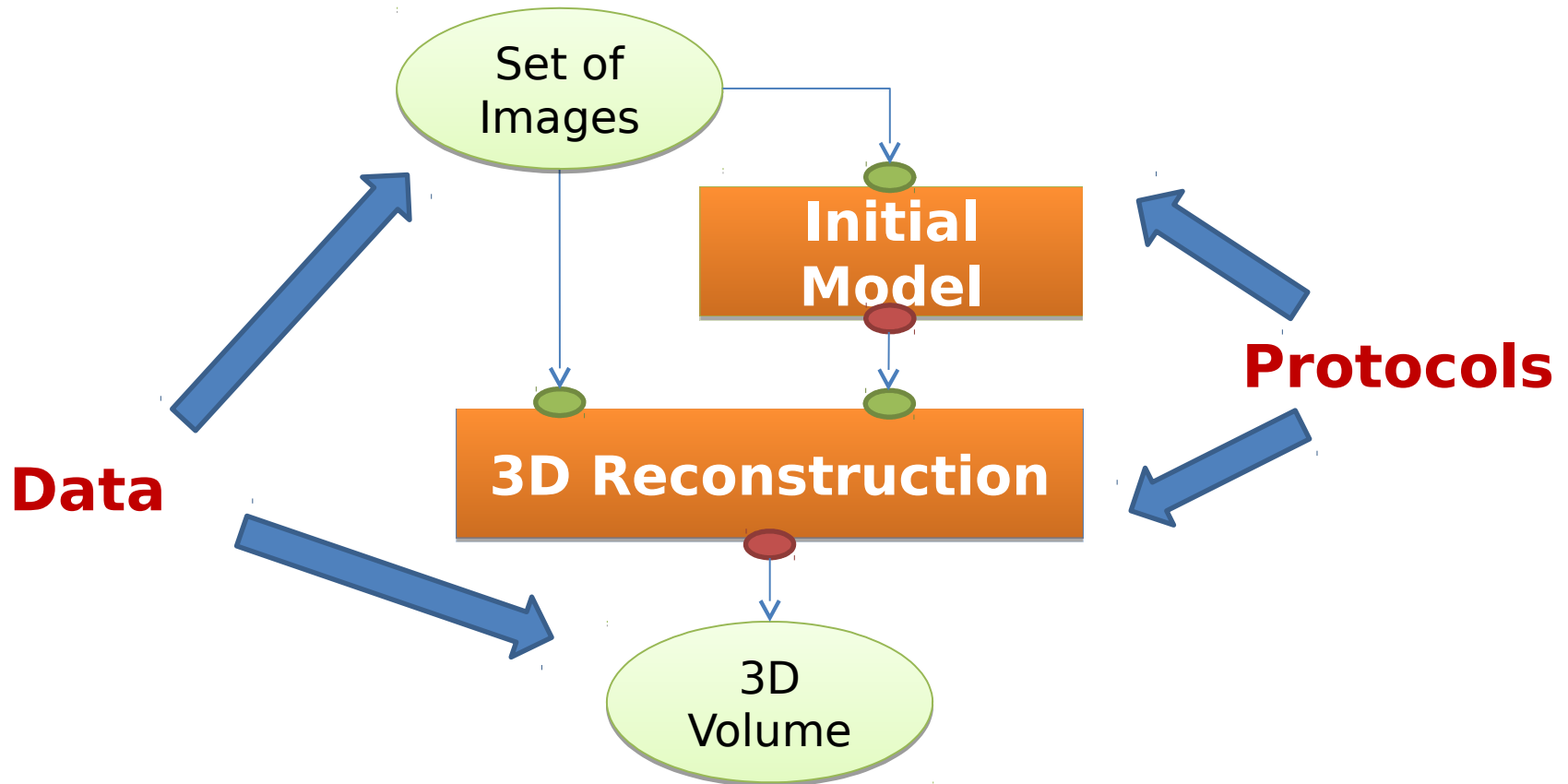
Using different EM software packages is now like the tower of Babel



# Scipion Principles

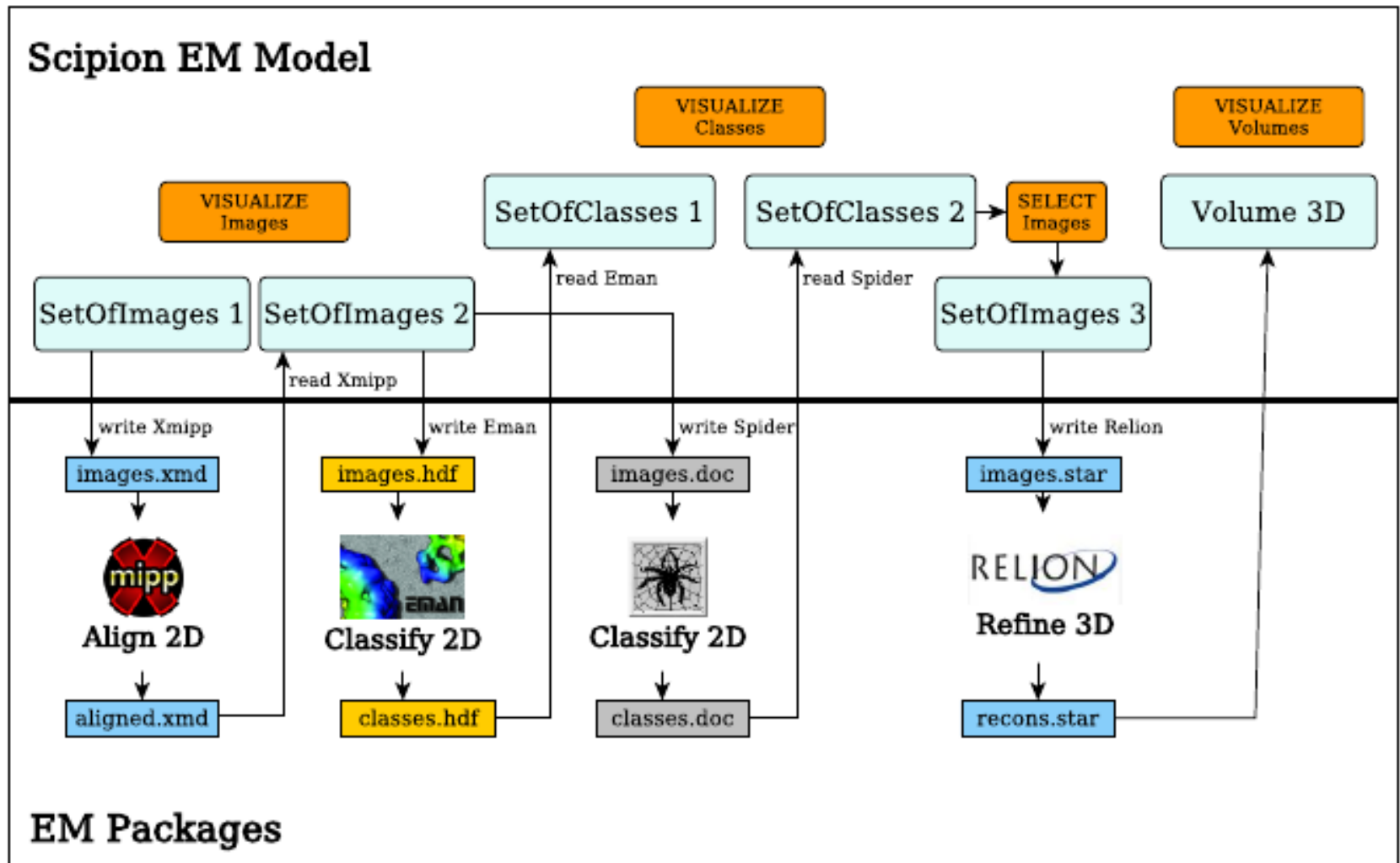
1. **Integrate** EM software packages to be used in the same project.
2. Full project traceability, improving **reproducibility**.
3. Execute complete workflows in an **automated** manner.
4. Easy to install and use.
5. Easy to extend with new protocols.

# We created a model of the EM domain

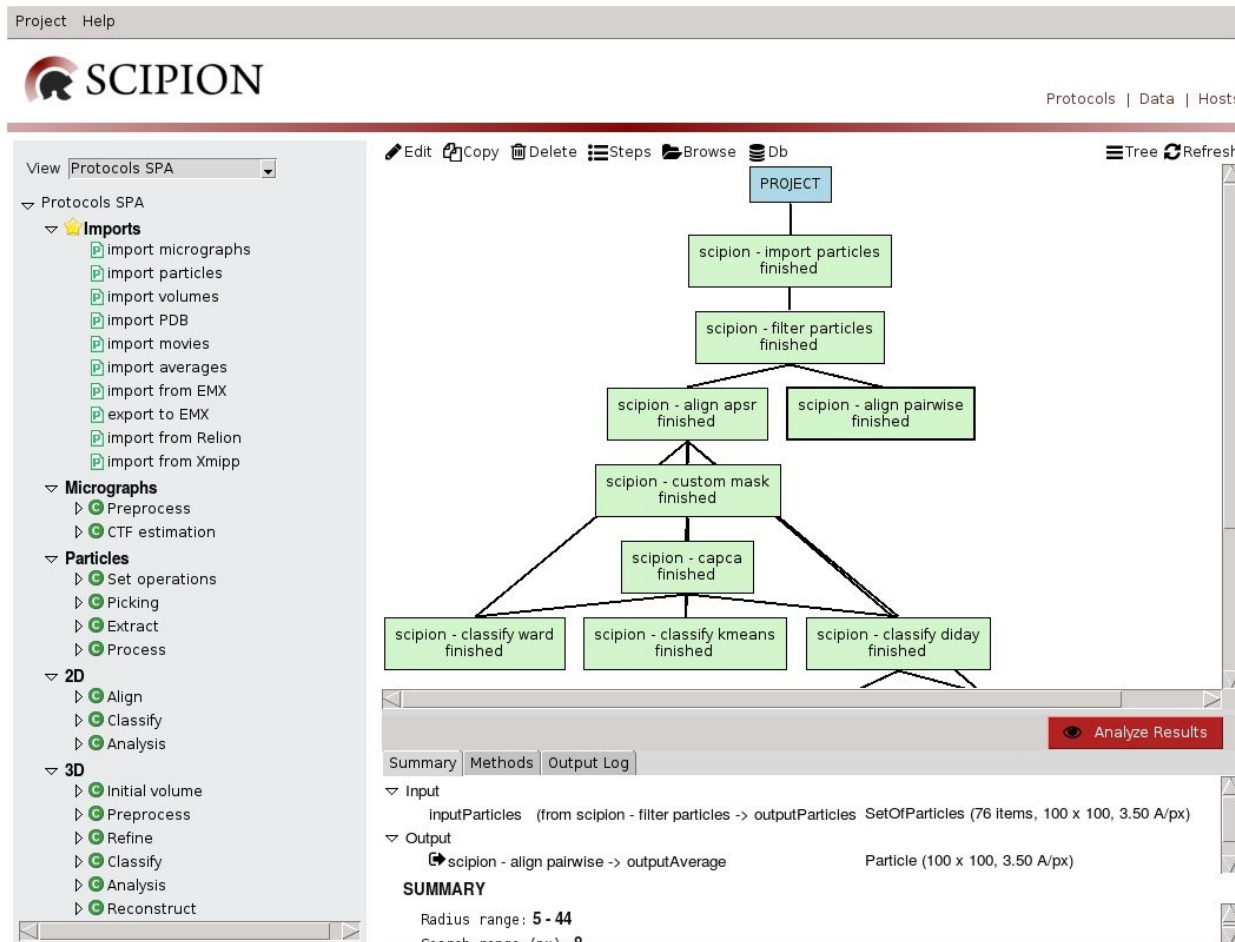




# Conversion functions bridge between packages



# We track all the steps performed in a project



# All parameters are also stored

**A** Protocol: spider - filter particles

Run

Run name: spider - filter particles Comment:

Run mode: ☒ Resume ☐ Restart Host: localhost

Parallel: Threads 2 MPI 1 Use queue? ☐ Yes ☒ No

Expert Level: ☒ Normal ☐ Advanced

Input

Input Particles: 3mics particles

Filter type: Butterworth

Filter mode: low-pass

Use padding? ☒ Yes ☐ No

Frequency: Highest 0.5

Close Save Execute

**B** Package References:

Sheikh, et al. Nature Protocols, 2008

Frank, et al. JSB, 1996

OK

**C**

Object	Info	Creation
3mics particles	[1762 items, 64 x 64, 3.54 A/pix]	2015-05-13 10:00:36
extract all mics->outputParticles	[6679 items, 64 x 64, 3.54 A/pix]	2015-05-13 10:35:08
removed z>3->outputParticles	[6362 items, 64 x 64, 3.54 A/pix]	2015-05-13 11:05:50
removed z>3->outputParticles	[6362 items, 64 x 64, 3.54 A/pix]	2015-05-13 11:05:51

Double click an item to preview the object

Select Cancel

**D**

Object

- 001@Falcon\_2012\_0
- 002@Falcon\_2012\_0
- 003@Falcon\_2012\_0
- 004@Falcon\_2012\_0
- 005@Falcon\_2012\_0
- 006@Falcon\_2012\_0
- 007@Falcon\_2012\_0
- 008@Falcon\_2012\_0
- 009@Falcon\_2012\_0
- 010@Falcon\_2012\_0
- 011@Falcon\_2012\_0
- 012@Falcon\_2012\_0
- 013@Falcon\_2012\_0

Particle

Filtered particle

Frequencies

Low freq: 0.10 High freq: 0.50 Preview

Select Cancel

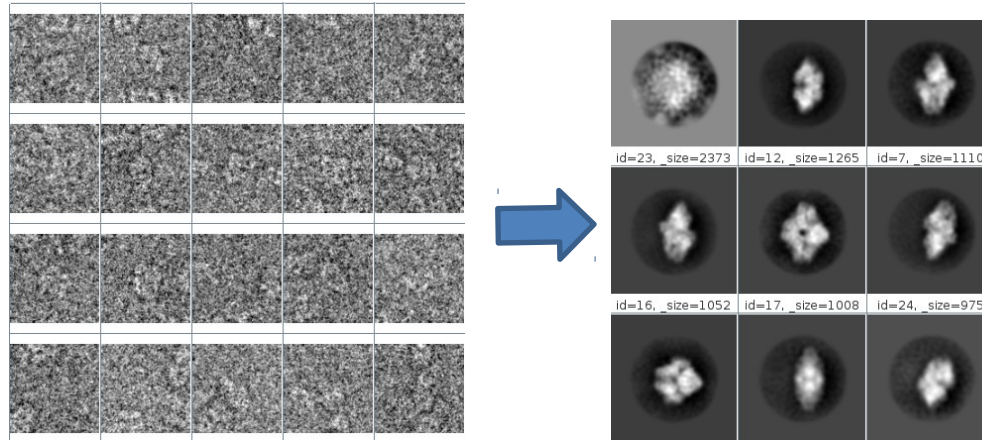
# Data Challenges: Accessibility, Transparency



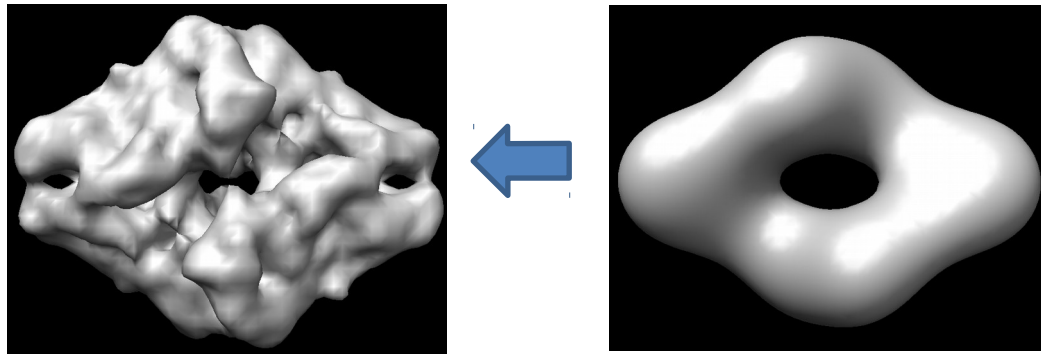


# Avoid data duplication as much as possible

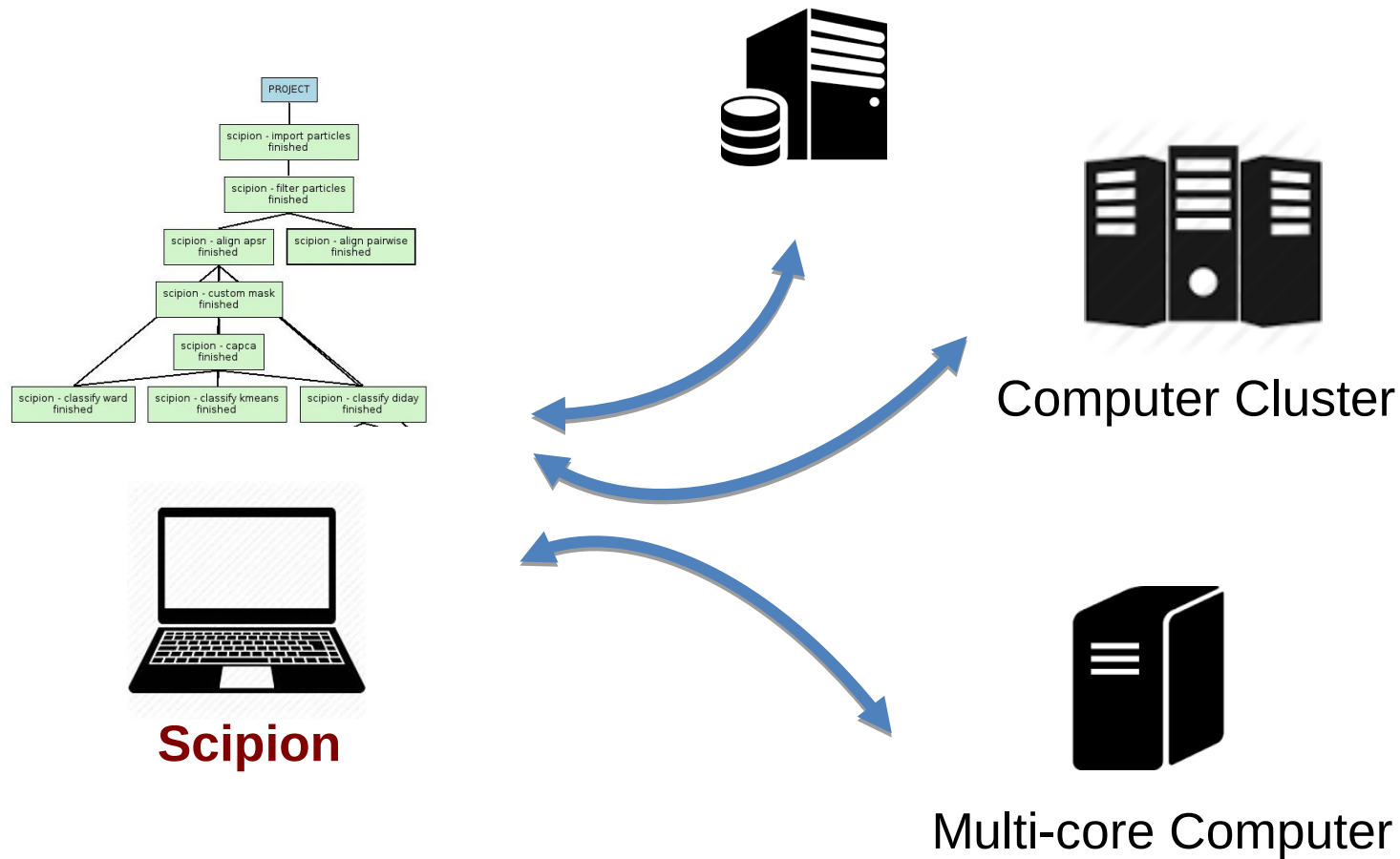
2D classification: Only store alignment parameters



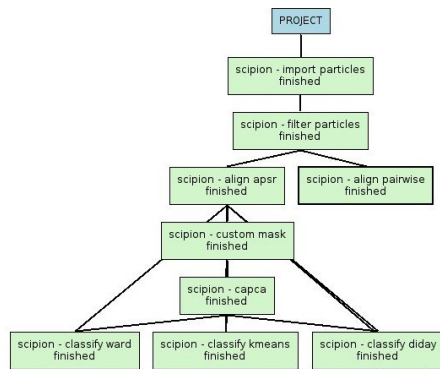
New 3D objects stored per iteration



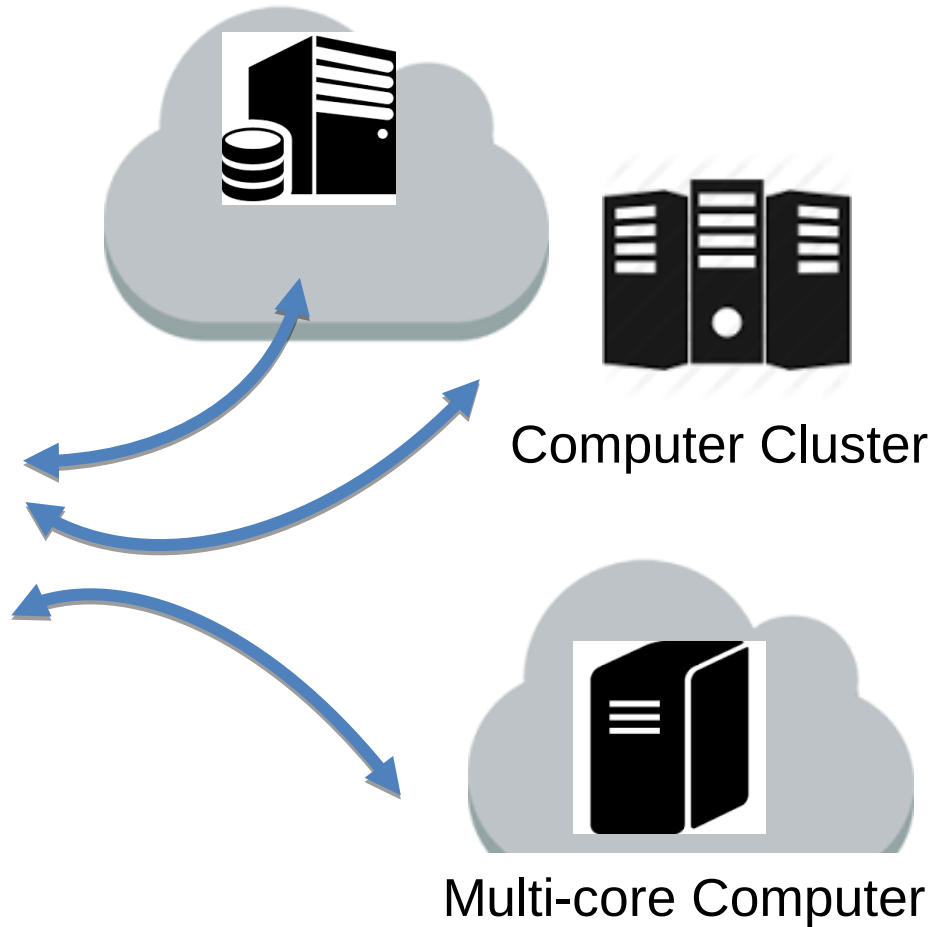
# Distribute the processing workflow, dealing with heterogeneous computing requirements



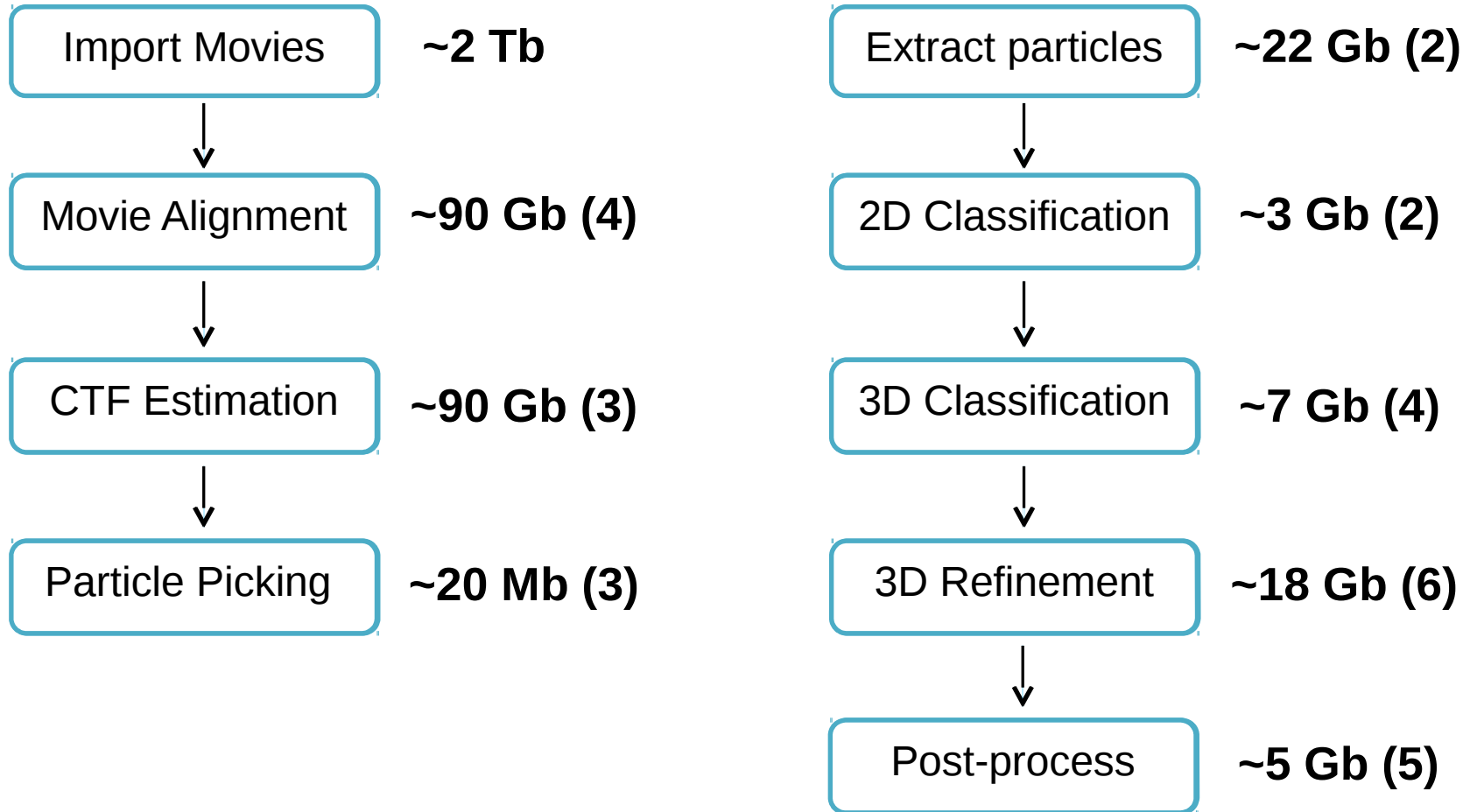
# Distribute the processing workflow, dealing with heterogeneous computing requirements



**Scipion**

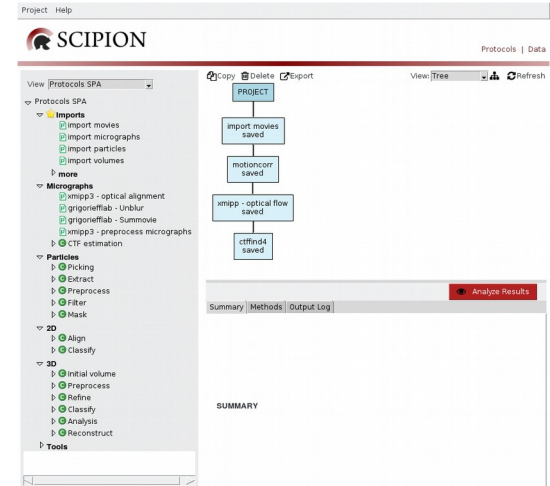
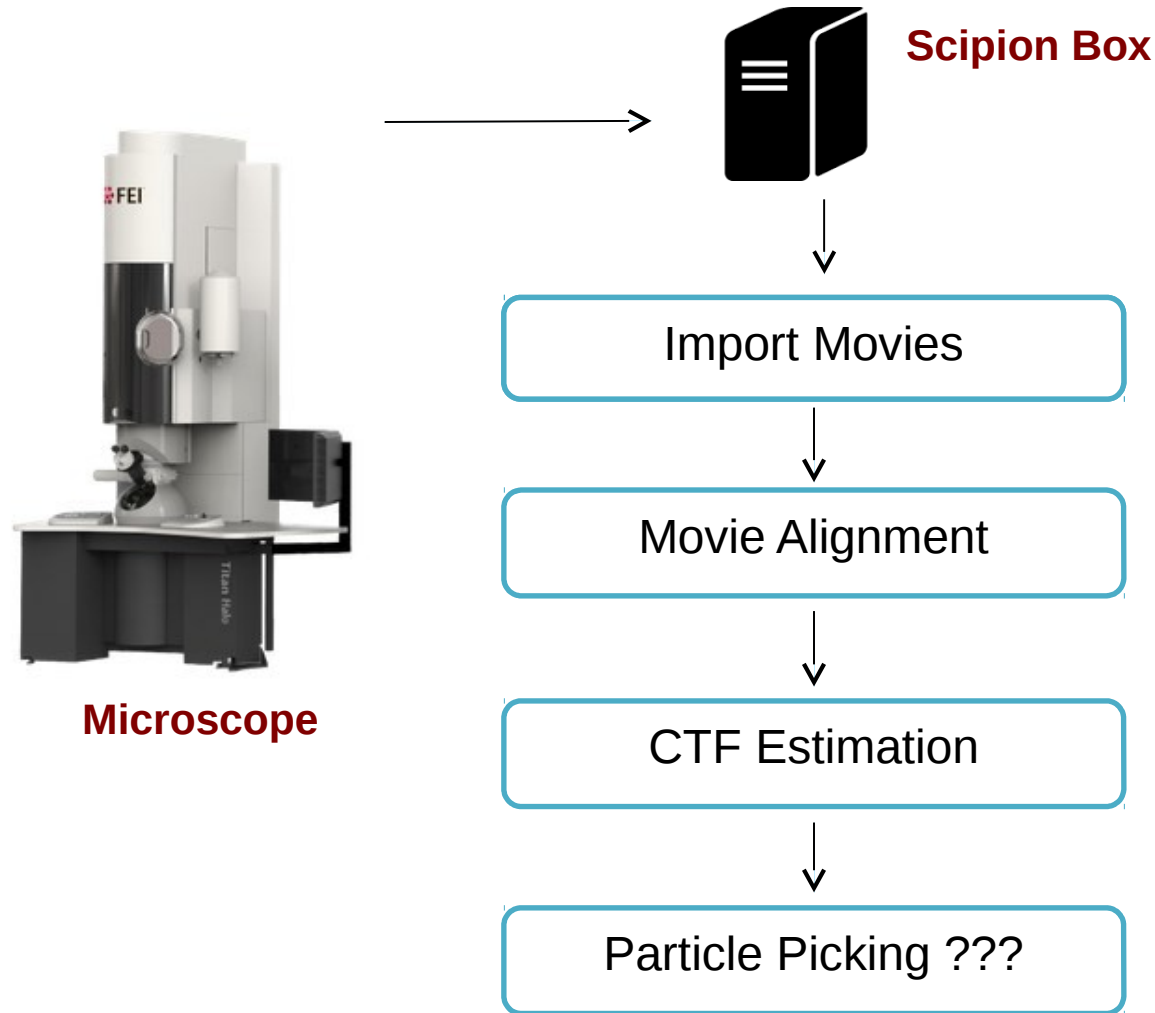


# Most of individual jobs involves large size input/output datasets

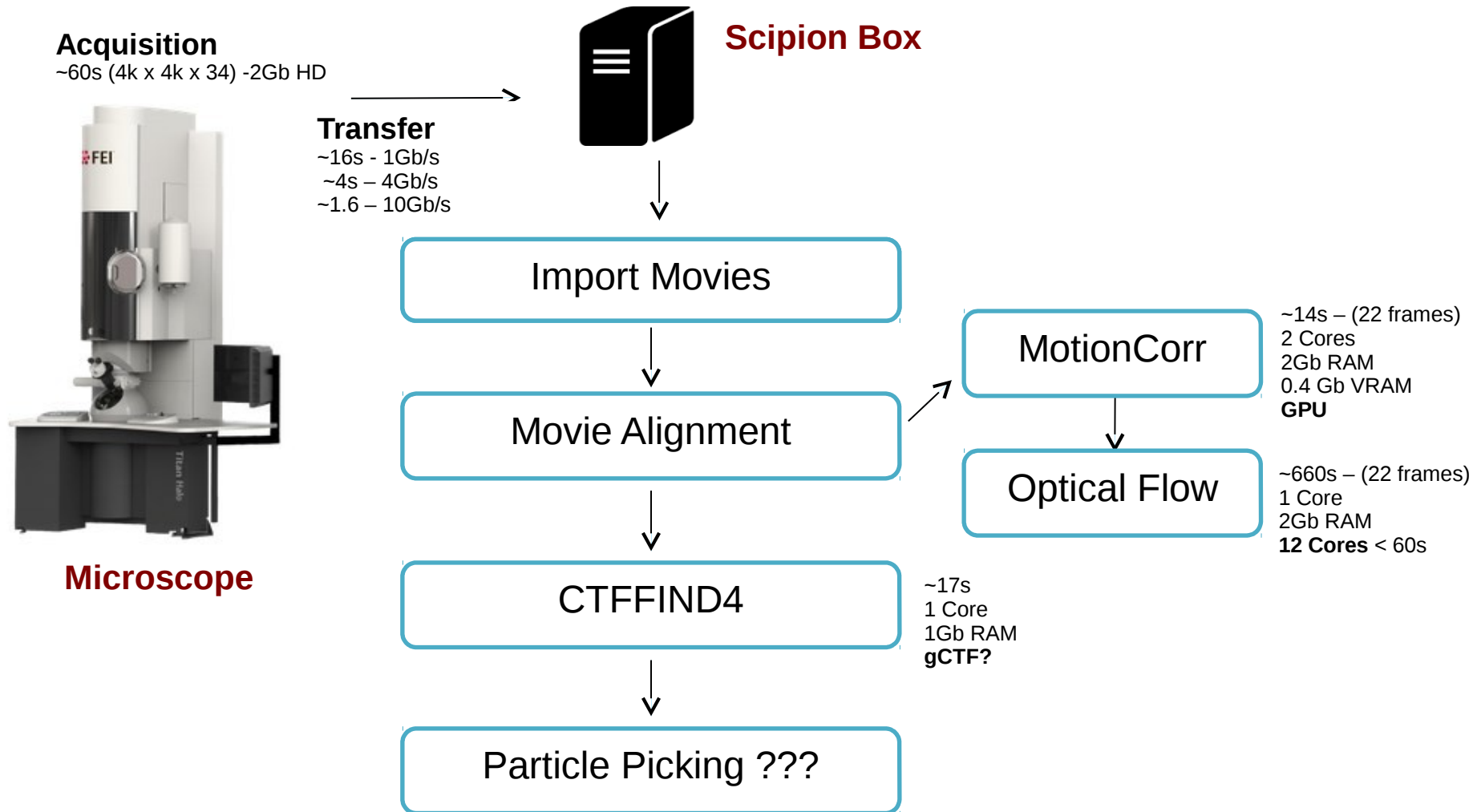




# Preprocess data close to the Source



# Run workflows automatically in streaming



# Scipion is open-source and freely available

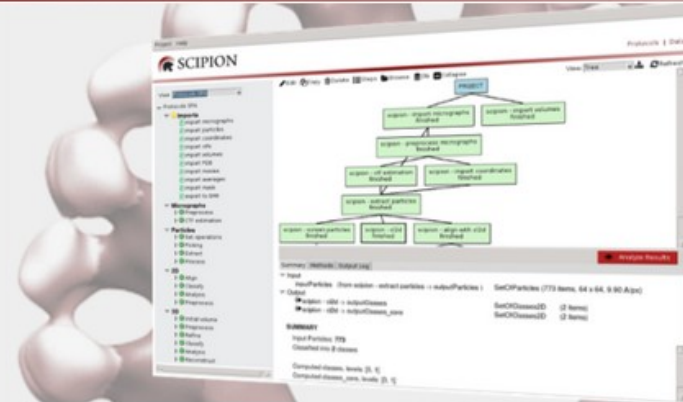
## <http://scipion.cnb.csic.es>



## Integration, Reproducibility and Analysis

Scipion is an image processing framework to obtain 3D models of macromolecular complexes using Electron Microscopy.

Download



## Documentation

Find out more information about Scipion for both users and developers. Check how to install Scipion, the list of integrated protocols and some introductory tutorials. Learn how to extend Scipion with new protocols, its API or description of our development tools.

## Web Tools

Give a try to our online processing workflows, which are a subset of the protocols in Scipion to provide a first try without any local installation. Current web-tools include initial volume estimation, movie alignment and local resolution (ResMap).

## News and Events

Keep an eye on latest Scipion news. Check out about bugfixes, new features and release plan. Don't miss the next Scipion workshop or any related event. Don't hesitate to contact us for any feedback or if you want to organize a workshop at your institution.

© 2015 - Scipion is licensed under the GNU General Public License



## Instruct Centres > Centre for Image Processing

### Instruct Image Processing Center (I2PC)



### Instruct Flagship Platform ☆

#### Personalized support for Image Processing project 🔬

The Instruct Image Processing Center provide personalized support to Instruct projects requiring specialized processing of either electron or X-ray microscopy images. Currently our main offer is focused in the field of single particles analysis using Xmipp and Scipion softwares. Xmipp is a comprehensive suite of image processing algorithms with a strong emphasis in the analysis of single particles, although it is extending towards electron and X-ray tomography. Scipion, in turn, provides a multi-package workflow-oriented platform bridging several software suites, like EMAN, Spider, Relion, Bsoft, FREALIGN and ctffind. More information can be found at: <http://xmipp.cnb.csic.es> and <http://scipion.cnb.csic.es/docs>

[Instruct platform information →](#)



# Conclusions (EM?)

1. We need **integrative** software frameworks to combine different algorithms
2. **Reproducibility** should be a first priority task to move science forward.
3. **Computing** requirements are very heterogeneous (hardware and software).
4. **Data Management** is a very important issue.
5. We need to **join efforts** v(user, developers and infrastructures) through collaborative environments.





[www.structuralbiology.eu](http://www.structuralbiology.eu)  
Follow us on twitter  
**@instructhub**

