netherlands eScience center

Automated electron diffraction using *Instamatic*

Stef Smeets
17/12/2020

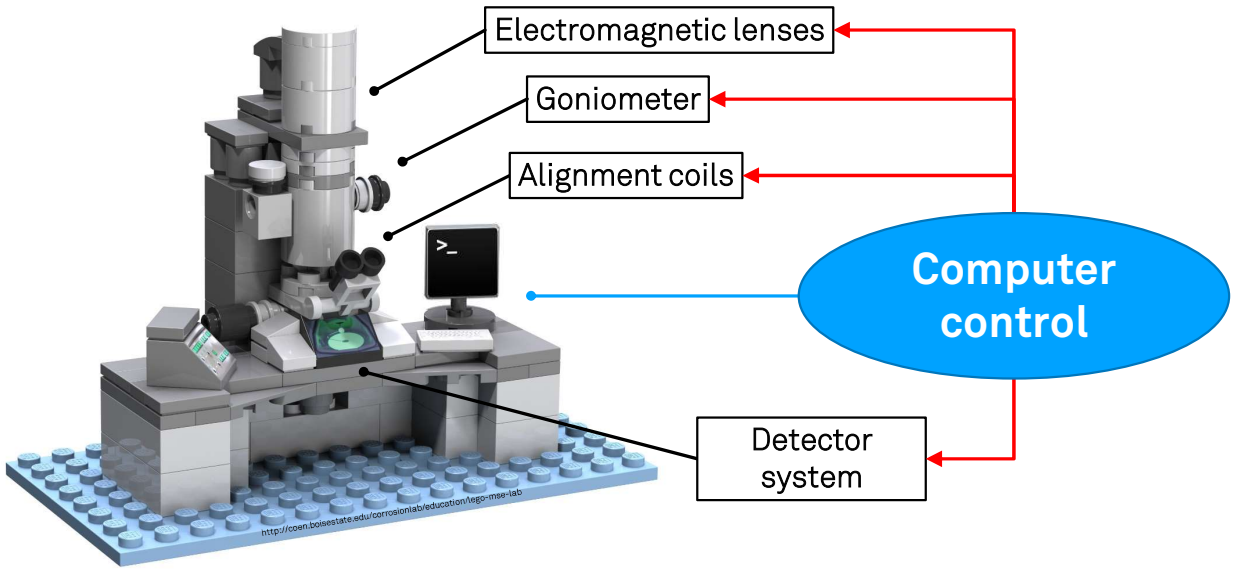
> *Compared to C/C++, **Python** can be used to write a much more **flexible** input mechanism in a much **shorter time**, so that **scientists** can feed the program its **data**, in all the **variations** that are required for reasons of **experimental setup**.*

Guido van Rossum (1998)
Python as a glue language
https://www.python.org/doc/essays/omg-darpa-mcc-position/

> *Python is ideal for oddball integration tasks.*

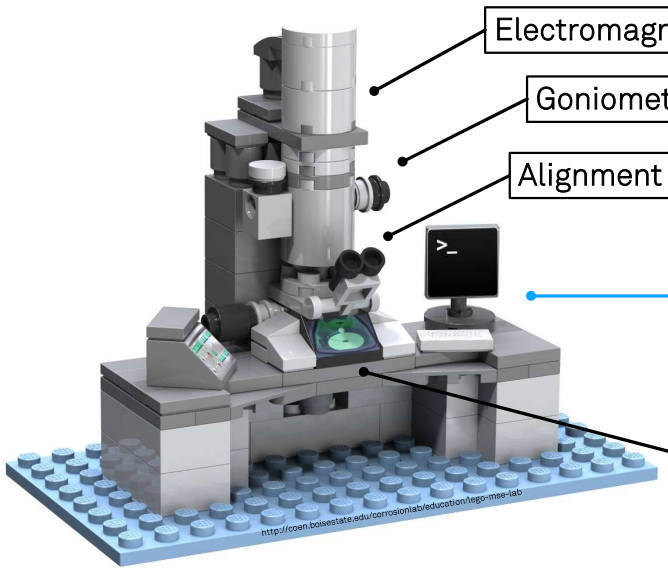
Guido van Rossum (1998)
Python as a glue language
https://www.python.org/doc/essays/omg-darpa-mcc-position/

Automation toolkit: *Instamatic*



Source code:
http://github.com/instamatic-dev/instamatic

Automation toolkit: *Instamatic*



Electromagnetic lenses

Goniometer

Alignment coils

Instamatic
(Python 3.6+)

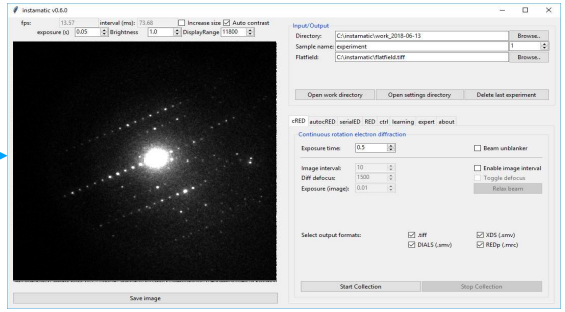
Detector system

Features

- Object oriented API
- Grid montaging
- Image segmentation
- Crystal tracking
- Automated data processing
- Calibrations/alignments

Automated experiments

- Screening
- High-throughput 3D ED
- Serial electron diffraction
- Unsupervised data collection



Microscopes

TFS Titan/Themis Z

JEOL 1400/2100/
3200/ARM200

Simulated

Cameras

ASI Cheetah

Gatan

Simulated

TVIPS (X)F416

Source code:
http://github.com/instamatic-dev/instamatic

When I started (~2014)

1. No open-source software
2. Manual crystal search
3. Semi-automated data collection
4. Not reproducible, ad-hoc protocols
5. Lack of experimental metadata
6. Time-consuming data processing

Instamatic (2017)

➡ http://github.com/instamatic-dev/instamatic

Automated crystal screening

Automated data collection with crystal tracking

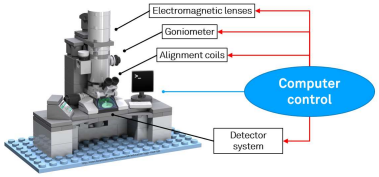
Reproducible experiments

Consistent metadata and logging

Data processing pipeline

When I started (~2014)

1. No open-source software
2. Manual crystal search
3. Semi-automated data collection
4. Not reproducible, ad-hoc protocols
5. Lack of experimental metadata
6. Time-consuming data processing



Instamatic (2017)

http://github.com/instamatic-dev/instamatic

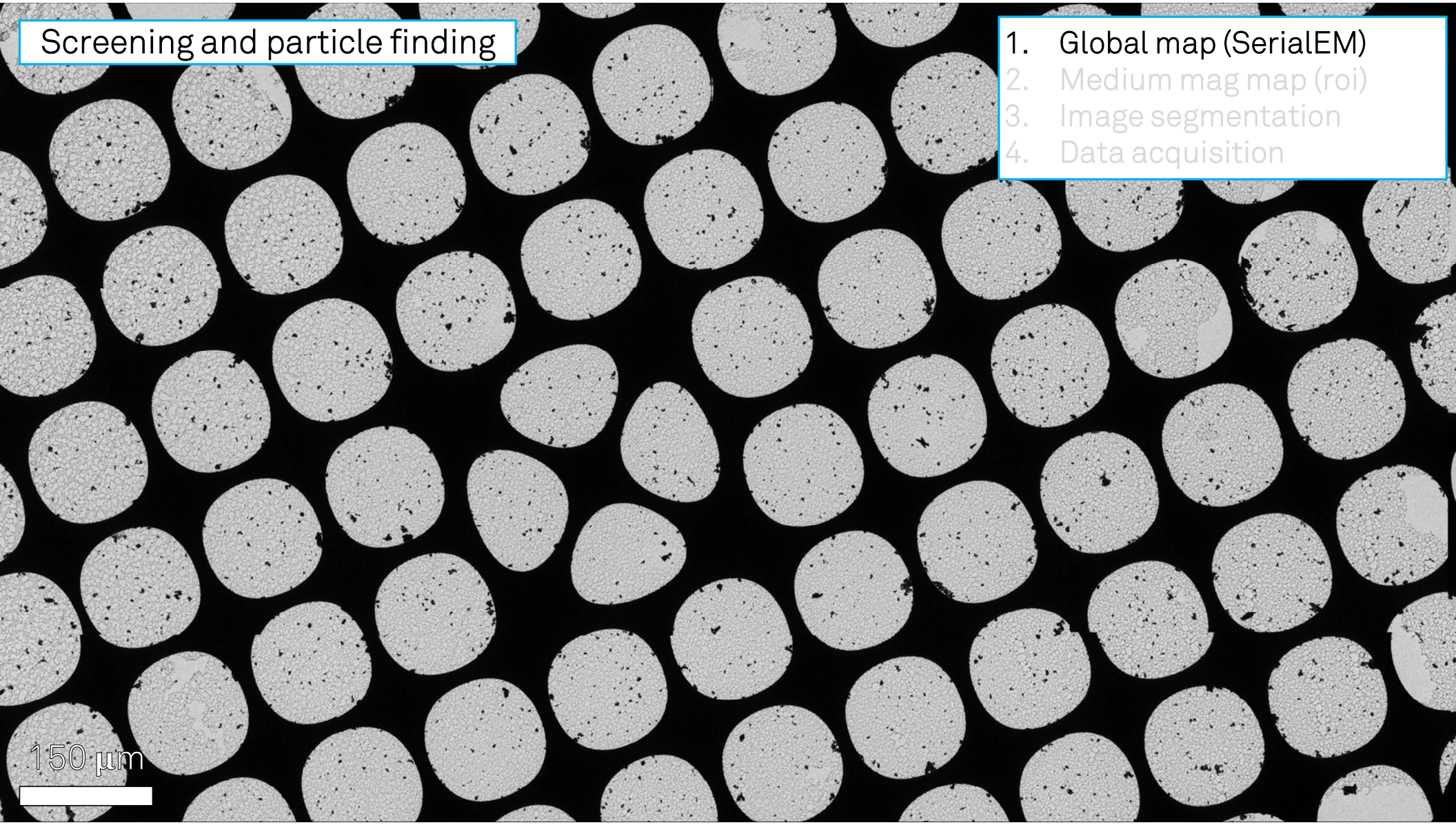
➡ Automated crystal screening

Automated data collection with crystal tracking

Reproducible experiments

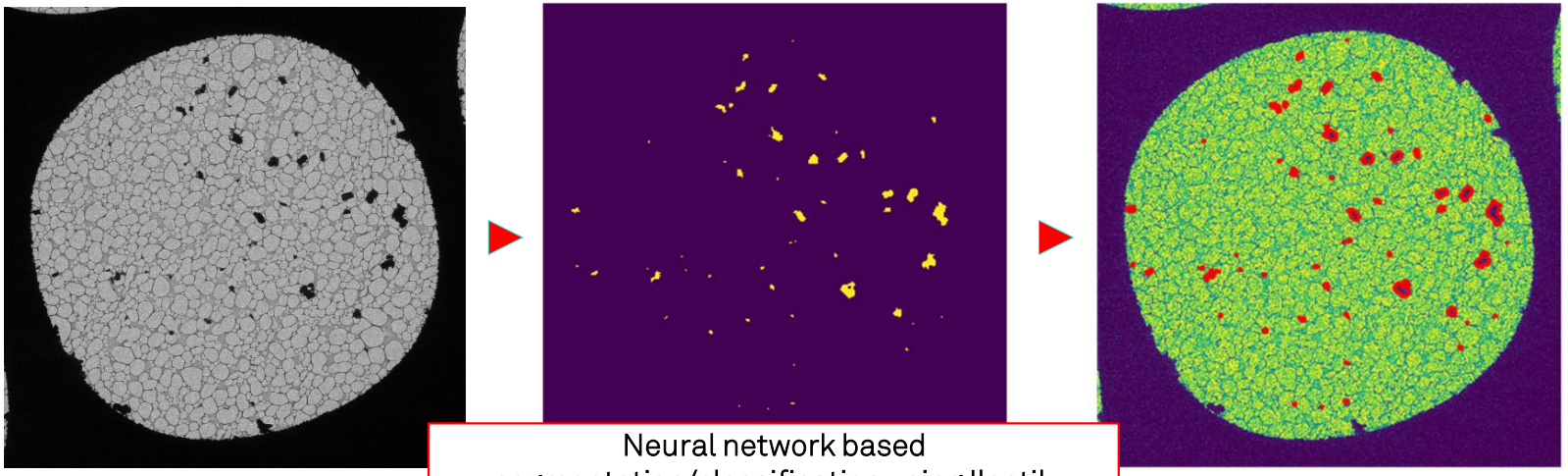
Consistent metadata and logging

Data processing pipeline


Screening and particle finding
1. Global map (SerialEM)
2. Medium mag map (roi)
3. Image segmentation
4. Data acquisition
150 µm

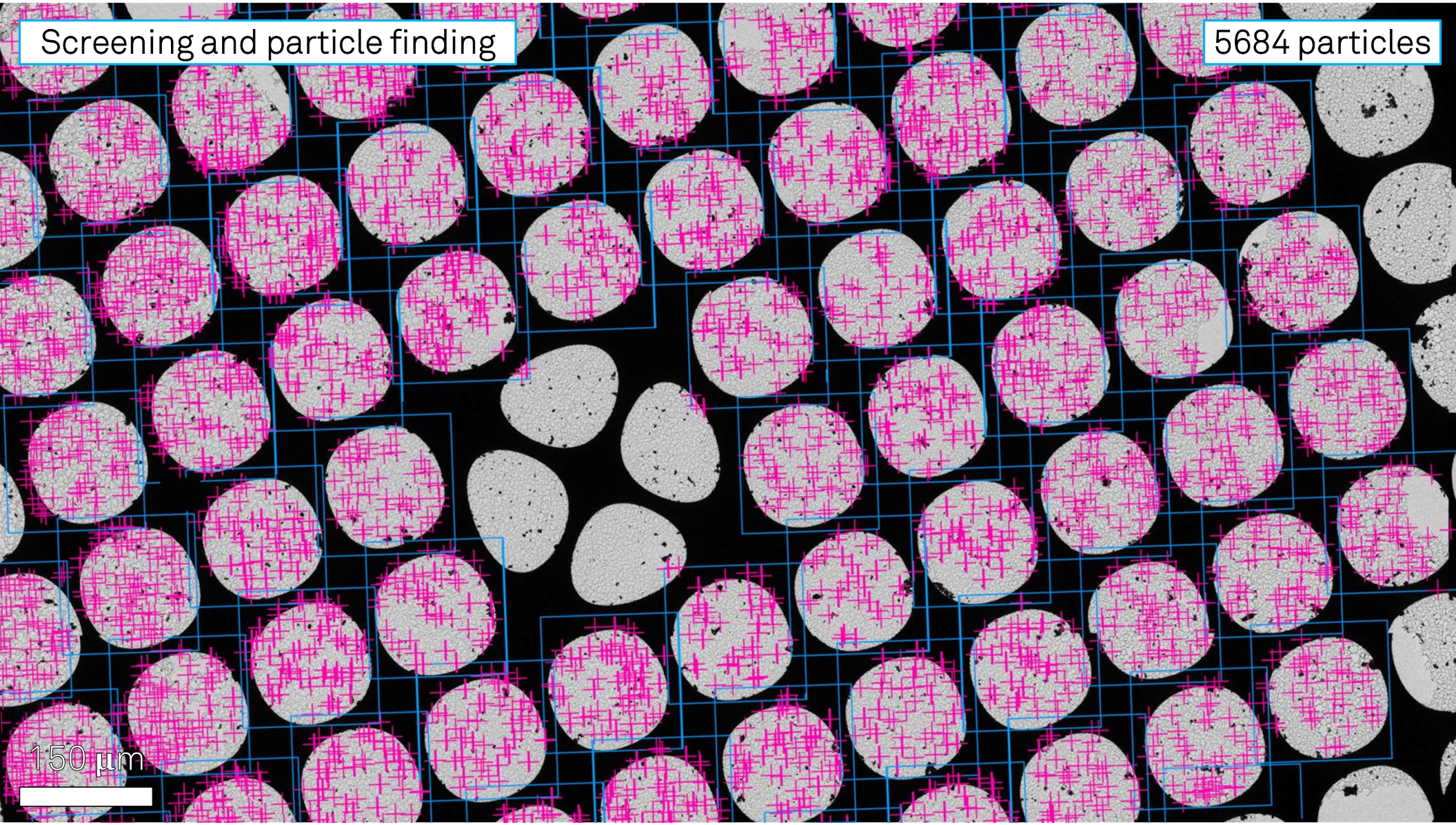
Screening and particle finding

1. Global map
2. Medium mag map (roi)
3. Image segmentation
4. Get particle coordinates



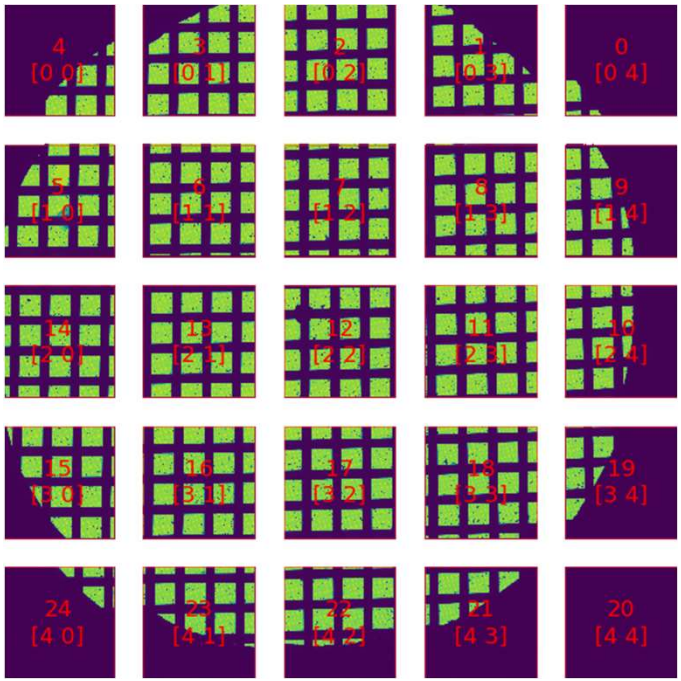
Neural network based
segmentation/classification using Ilastik
(https://gitlab.tudelft.nl/aj-lab/predicrystal)

150 µm

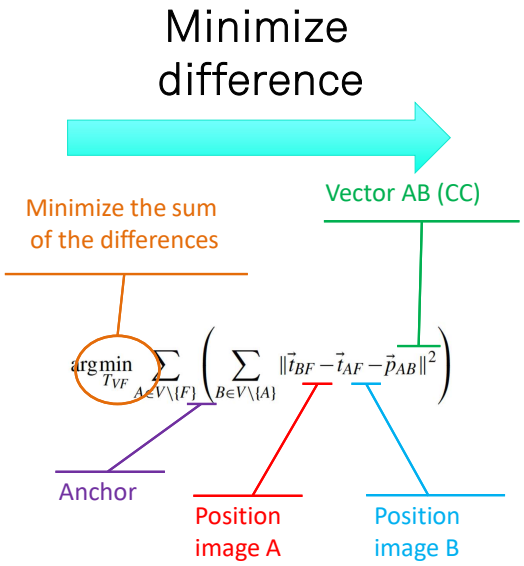

Screening and particle finding
5684 particles
150 µm

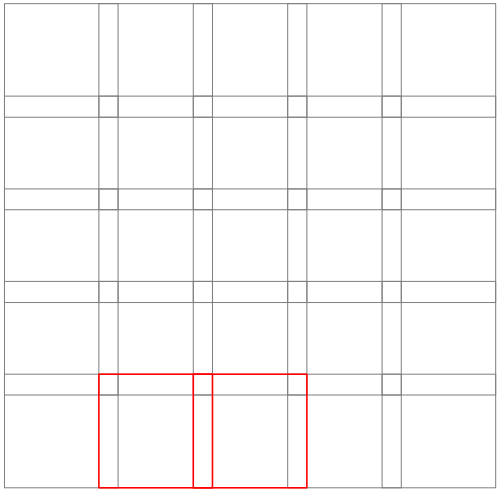
Montaging/stitching in *Instamatic*

Algorithm: Preibisch et al. (2009), Bioinformatics, 25(11):1463-1465
Implementation: https://github.com/instamatic-dev/pyserialem



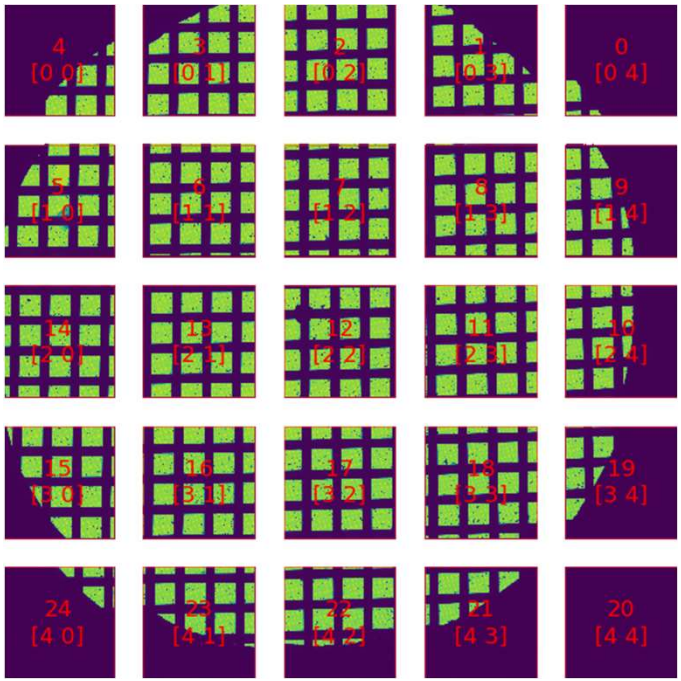
5x5 grid collected at low mag
10% overlap





Montaging/stitching in *Instamatic*

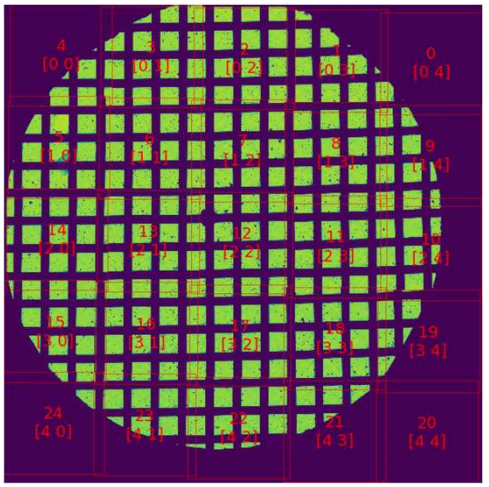
Algorithm: Preibisch et al. (2009), Bioinformatics, 25(11):1463-1465
Implementation: https://github.com/instamatic-dev/pyserialem



5x5 grid collected at low mag
10% overlap

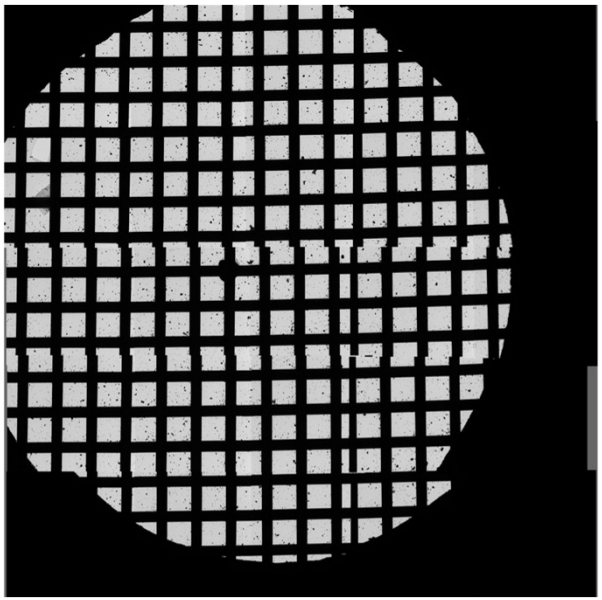
Minimize difference

$$\underset{T_{VF}}{\arg\min} \sum_{A \in V \setminus \{F\}} \left(\sum_{B \in V \setminus \{A\}} \|\vec{t}_{BF} - \vec{t}_{AF} - \vec{p}_{AB}\|^2 \right)$$

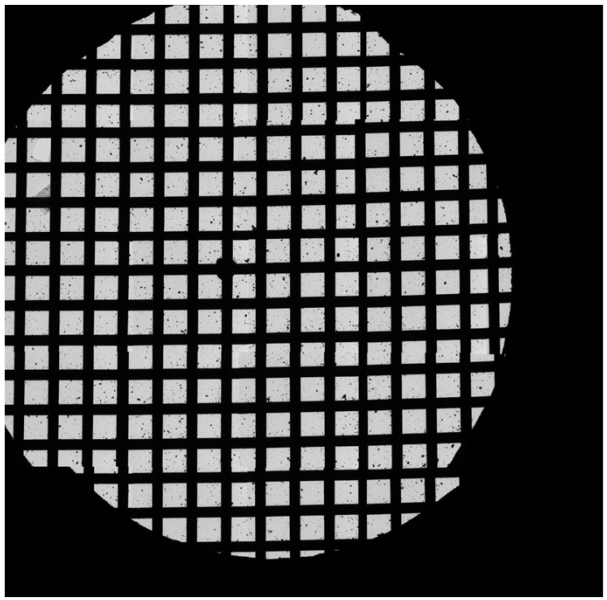


Stitched image

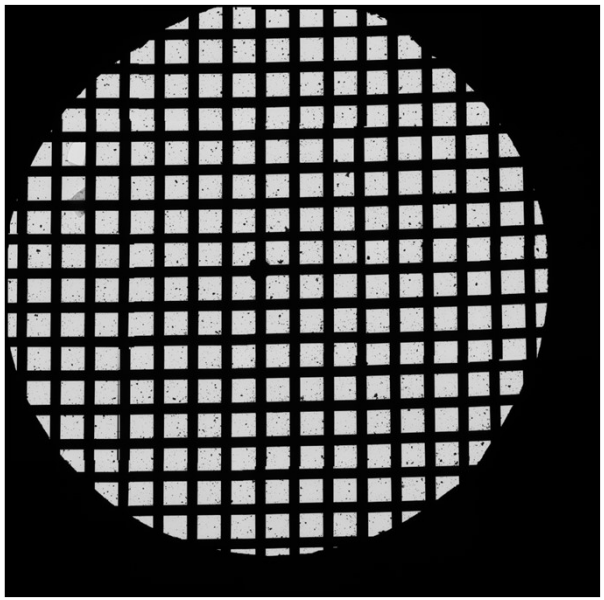
Stitching comparison



Data + Stitching
SerialEM



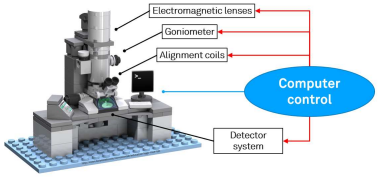
Data SerialEM
Stitching Instamatic

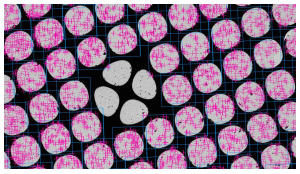


Data + stitching
Instamatic

When I started (~2014)

1. No open-source software
2. Manual crystal search
3. Semi-automated data collection
4. Not reproducible, ad-hoc protocols
5. Lack of experimental metadata
6. Time-consuming data processing





Instamatic (2017)

http://github.com/instamatic-dev/instamatic

Automated crystal screening

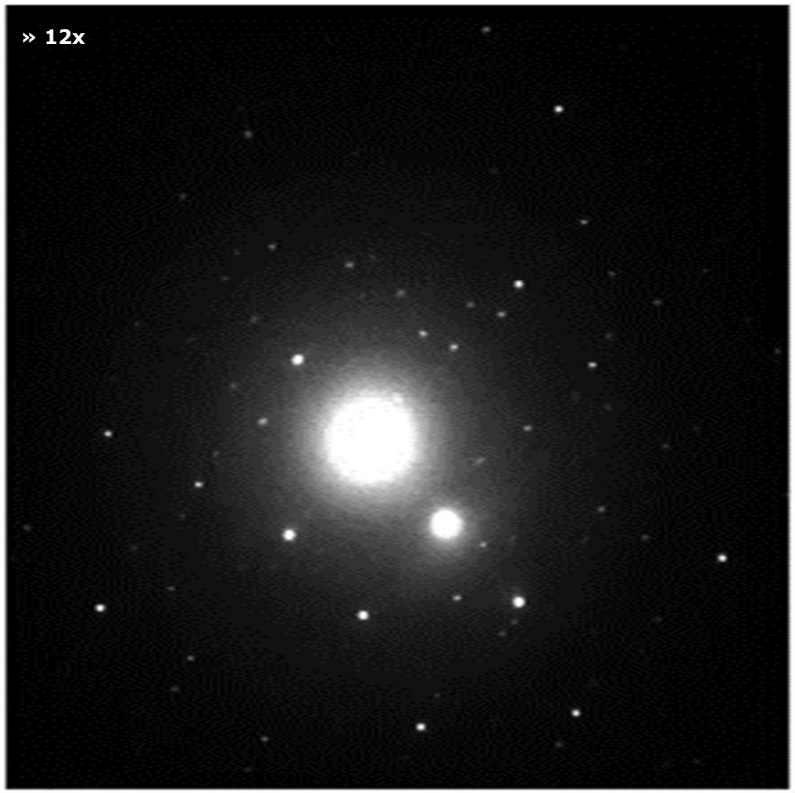
Automated data collection with crystal tracking

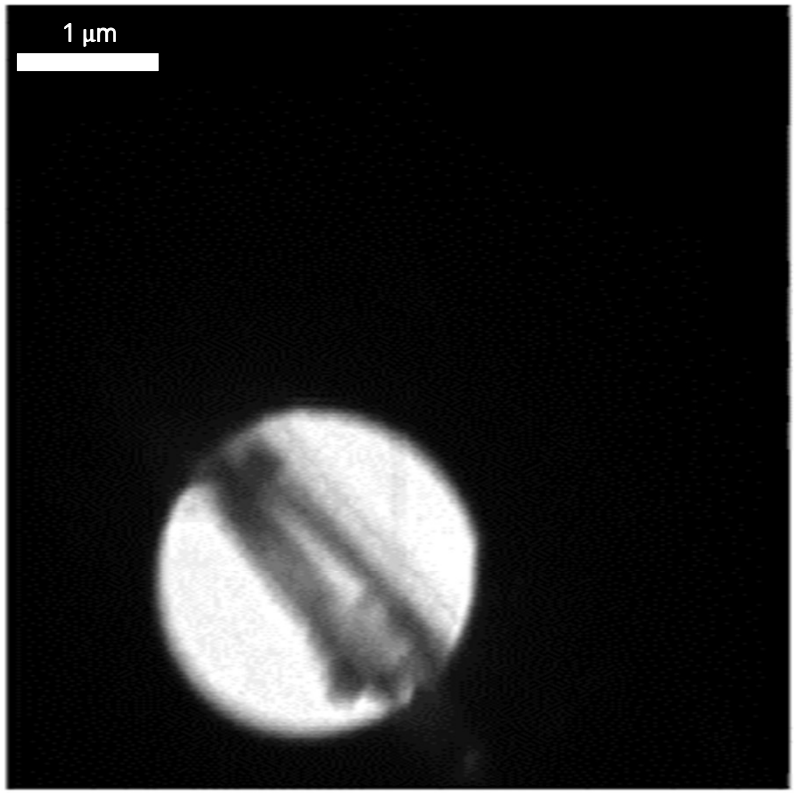
Reproducible experiments

Consistent metadata and logging

Data processing pipeline

Automated crystal tracking during continuous rotation

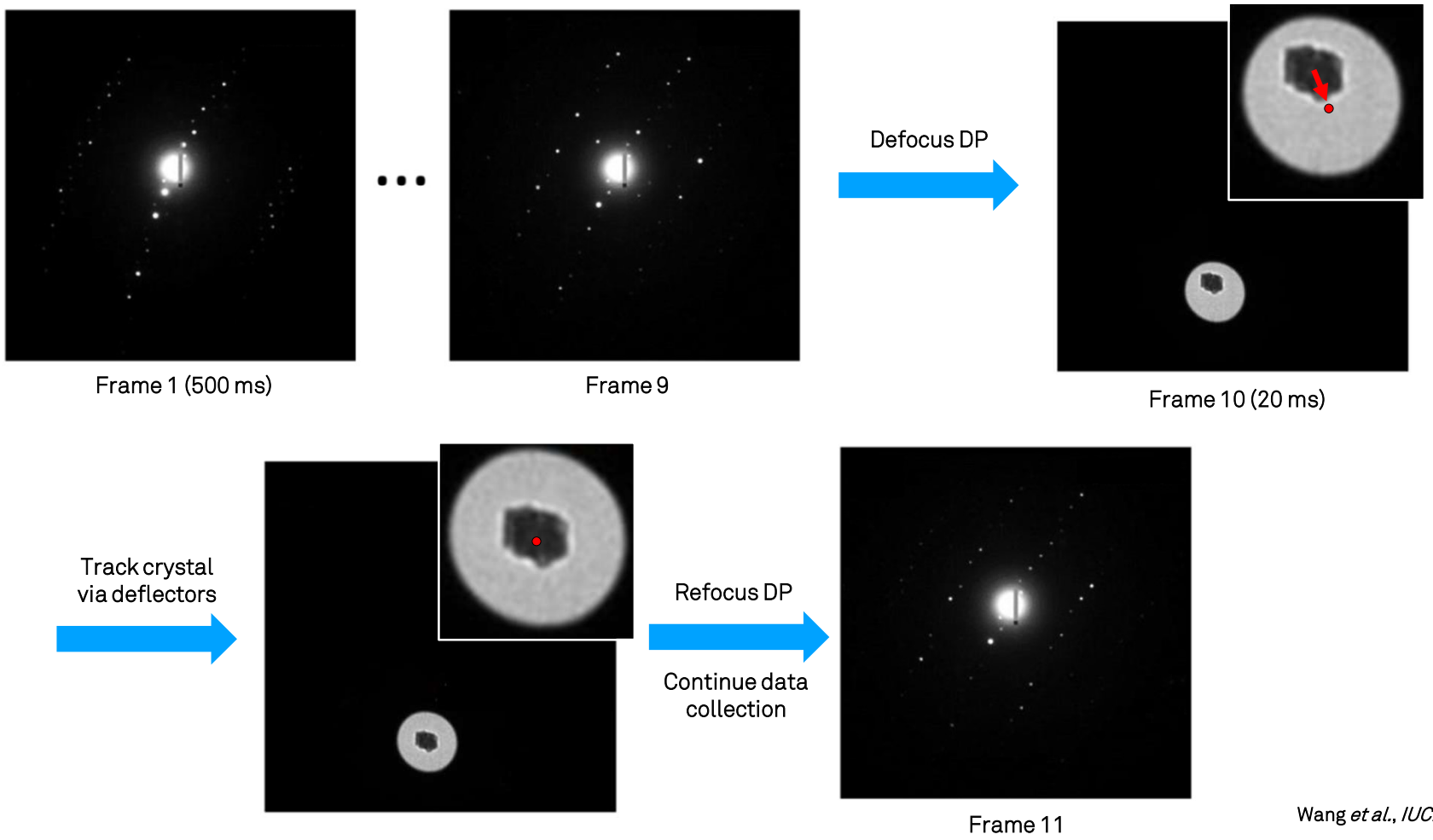




JEOL 2100-LaB_6 @ 200 kV (Timepix)
Rotation: -44.0 to 47.4° @ 0.76°/s (91.4°)
Exposure: 0.5 s, oscillation angle: 0.39°

Wang *et al.*, *IUCrJ* 6 (2019), 854-867

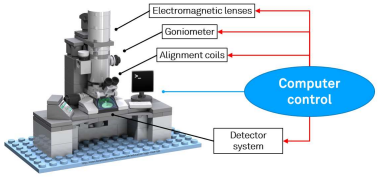
Crystal tracking strategy

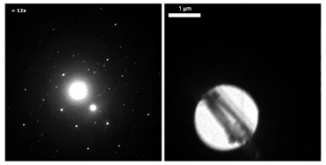


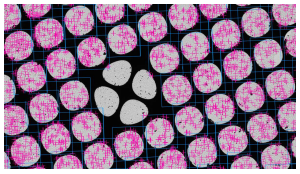
Wang *et al.*, *IUCrJ* 6 (2019), 854-867

When I started (~2014)

1. No open-source software
2. Manual crystal search
3. Semi-automated data collection
4. Not reproducible, ad-hoc protocols
5. Lack of experimental metadata
6. Time-consuming data processing





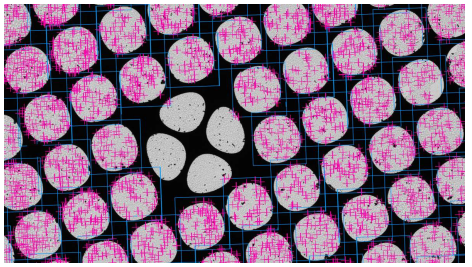


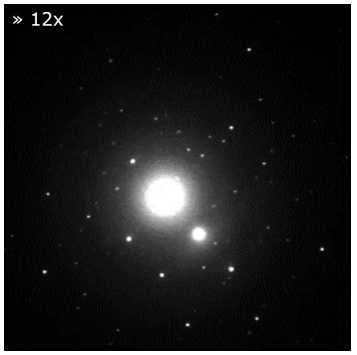
Instamatic (2017)

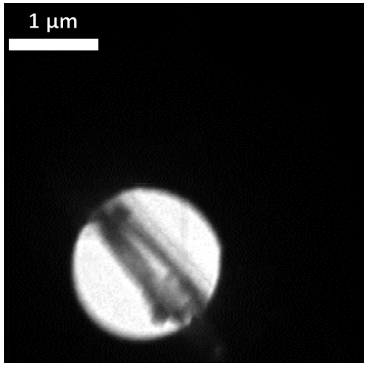
http://github.com/instamatic-dev/instamatic
Automated crystal screening
Automated data collection with crystal tracking
Reproducible experiments
Consistent metadata and logging
Data processing pipeline

Reproducible experiments

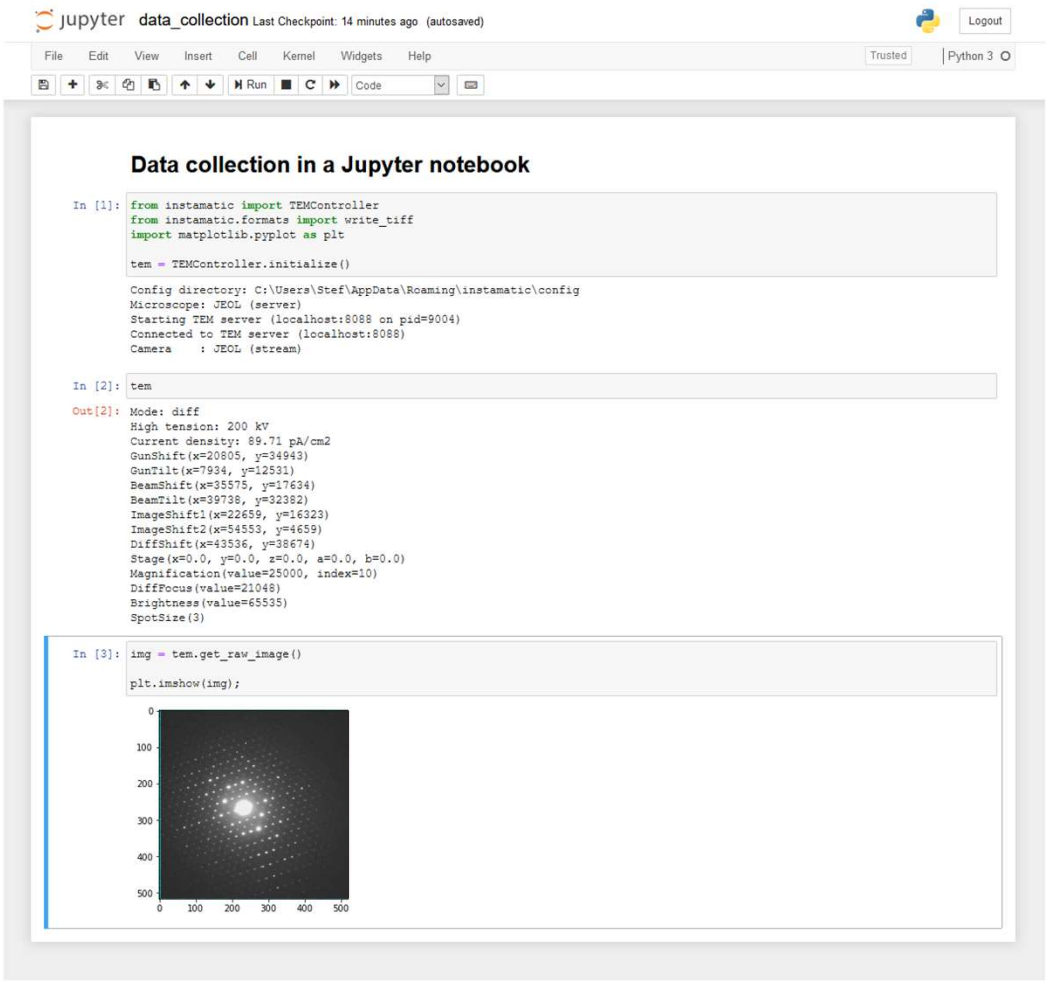
```python
from instamatic import TEMController
from instamatic.formats import read_tiff
from pyserialem import read_nav_file

ctrl = TEMController.initialize()
ctrl.mode.set('diff')

markers = read_nav_file('nav.nav', acquire_only=True)

def acquire_data(ctrl, tag):
    img, h = ctrl.get_image(exposure=0.2)
    write_tiff(f'data-{marker.tag}.tiff', img, header=h)

    ctrl.difffocus.defocus(offset=1500)

    img, h = ctrl.get_image(exposure=0.2)
    write_tiff(f'image-{marker.tag}.tiff', img, header=h)

    ctrl.difffocus.refocus()

for marker in markers:
    ctrl.stage.set(x=marker.x, y=marker.y)
    acquire_data(ctrl, tag)
```





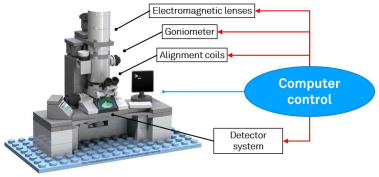


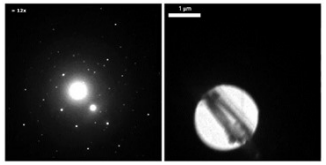
Instamatic in a Jupyter Notebook

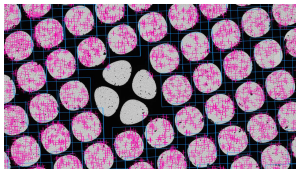


When I started (~2014)

1. No open-source software
2. Manual crystal search
3. Semi-automated data collection
4. Not reproducible, ad-hoc protocols
5. Lack of experimental metadata
6. Time-consuming data processing

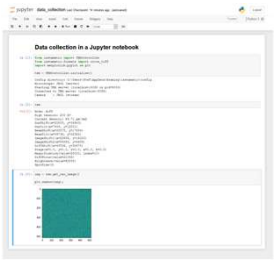






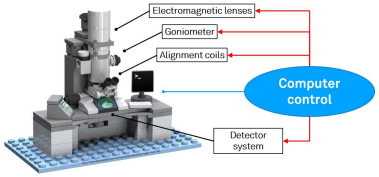
Instamatic (2017)

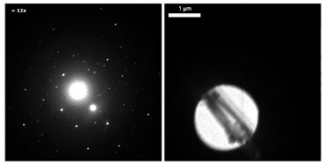
http://github.com/instamatic-dev/instamatic
Automated crystal screening
Automated data collection with crystal tracking
Reproducible experiments
Consistent metadata and logging
Data processing pipeline

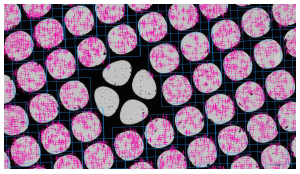


When I started (~2014)

1. No open-source software
2. Manual crystal search
3. Semi-automated data collection
4. Not reproducible, ad-hoc protocols
5. Lack of experimental metadata
6. Time-consuming data processing

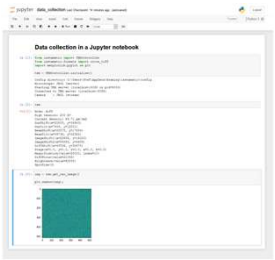


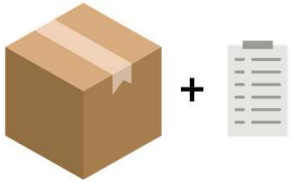




Instamatic

http://github.com/instamatic-dev/instamatic
Automated crystal screening
Automated data collection with crystal tracking
Reproducible experiments
Consistent metadata and logging
Data processing pipeline





Python data reduction pipeline

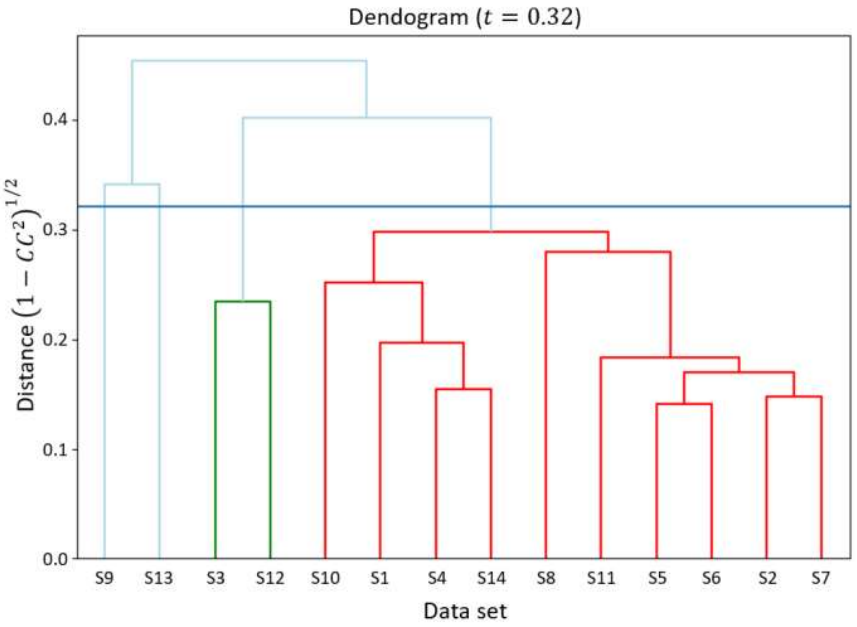
https://github.com/instamatic-dev/edtools

⬅ Data + metadata

1. Data conversion (image format, software input)
2. Data reduction (XDS)
3. Data scaling (XSCALE)
4. Cluster analysis (filter outliers, grouping)
5. Data merging (XSCALE)
6. Structure determination + refinement (SHELXS/SHELXL)

➔ Atomic model (higher precision)

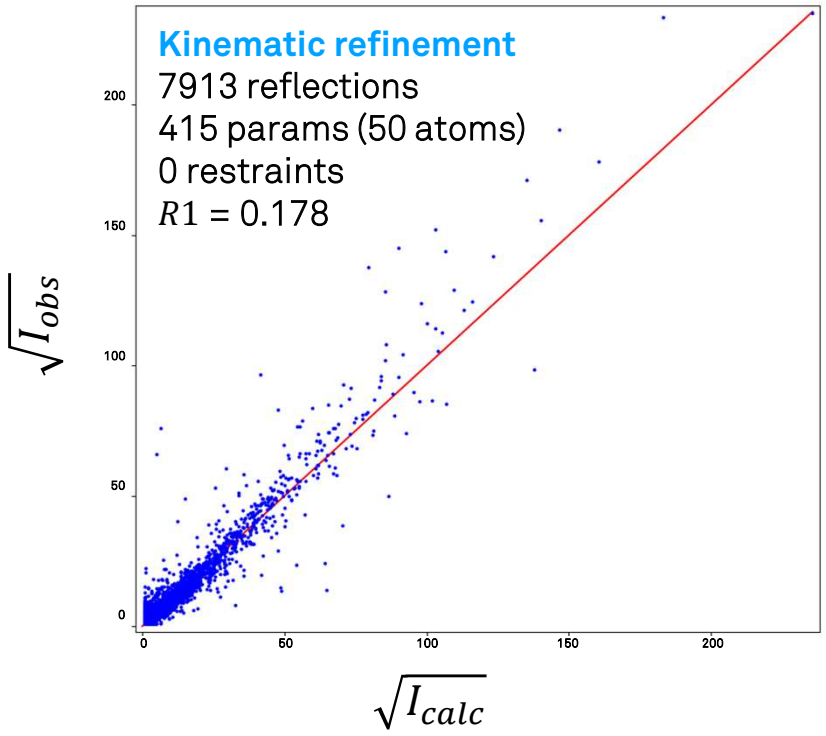
Hierarchical cluster analysis with SSZ-27 (14 crystals)

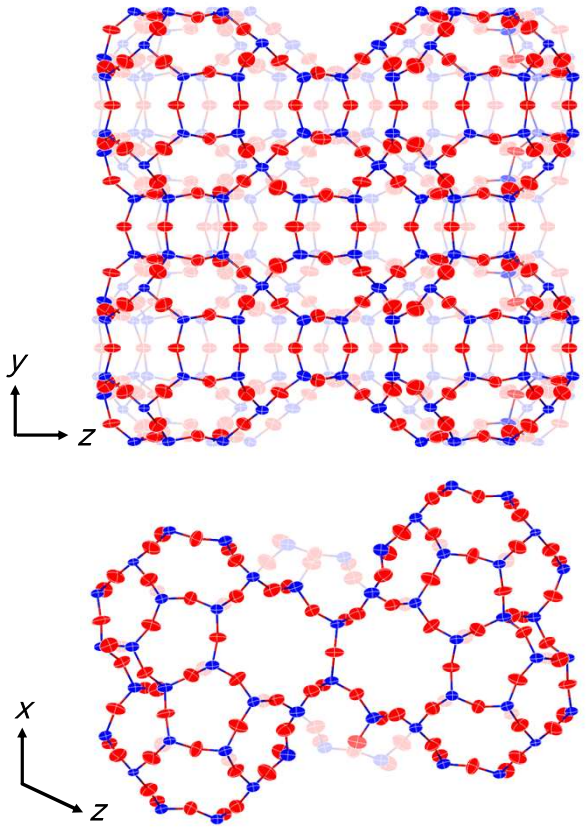


Shell	Full range	Low angle	High angle
Resolution range (Å)	11.59-0.80	11.59-1.0	1.0-0.80
Total data	130757	84285	46472
Unique Data	7914	4124	3790
Completeness (%)	98.8	99.6	97.8
$I/\sigma(I)$	4.53	7.01	1.82
R_{meas} (%)	31.5	29.3	141.3
$CC_{1/2}$	98.8	98.8	81.1
Overall B factor (Å^2)	8.17	-	-

Smeets *et al.*, *Angew. Chem. 58* (2019), 1380

Structure of SSZ-27 from 10 crystals





$[Si_{116}O_{232}]$

$\boldsymbol{C2/m}$

$a = 23.3$ Å

$c = 13.4$ Å

$c = 24.4$ Å

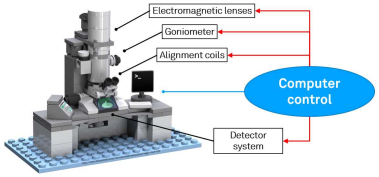
$\beta = 114.2°$

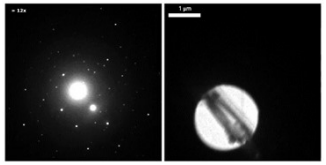
Smeets *et al.*, *Angew. Chem. 58* (2019), 1380

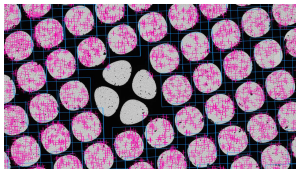
Instamatic

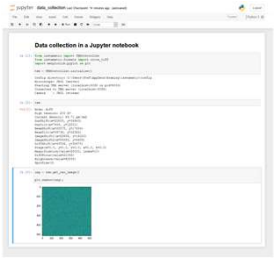
http://github.com/instamatic-dev/instamatic
Automated crystal screening
Automated data collection with crystal tracking
Reproducible experiments
Consistent metadata and logging
Data processing pipeline

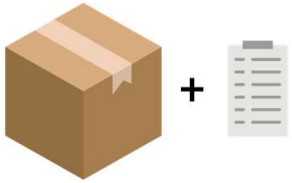
Automated online structure determination pipeline in *Instamatic*



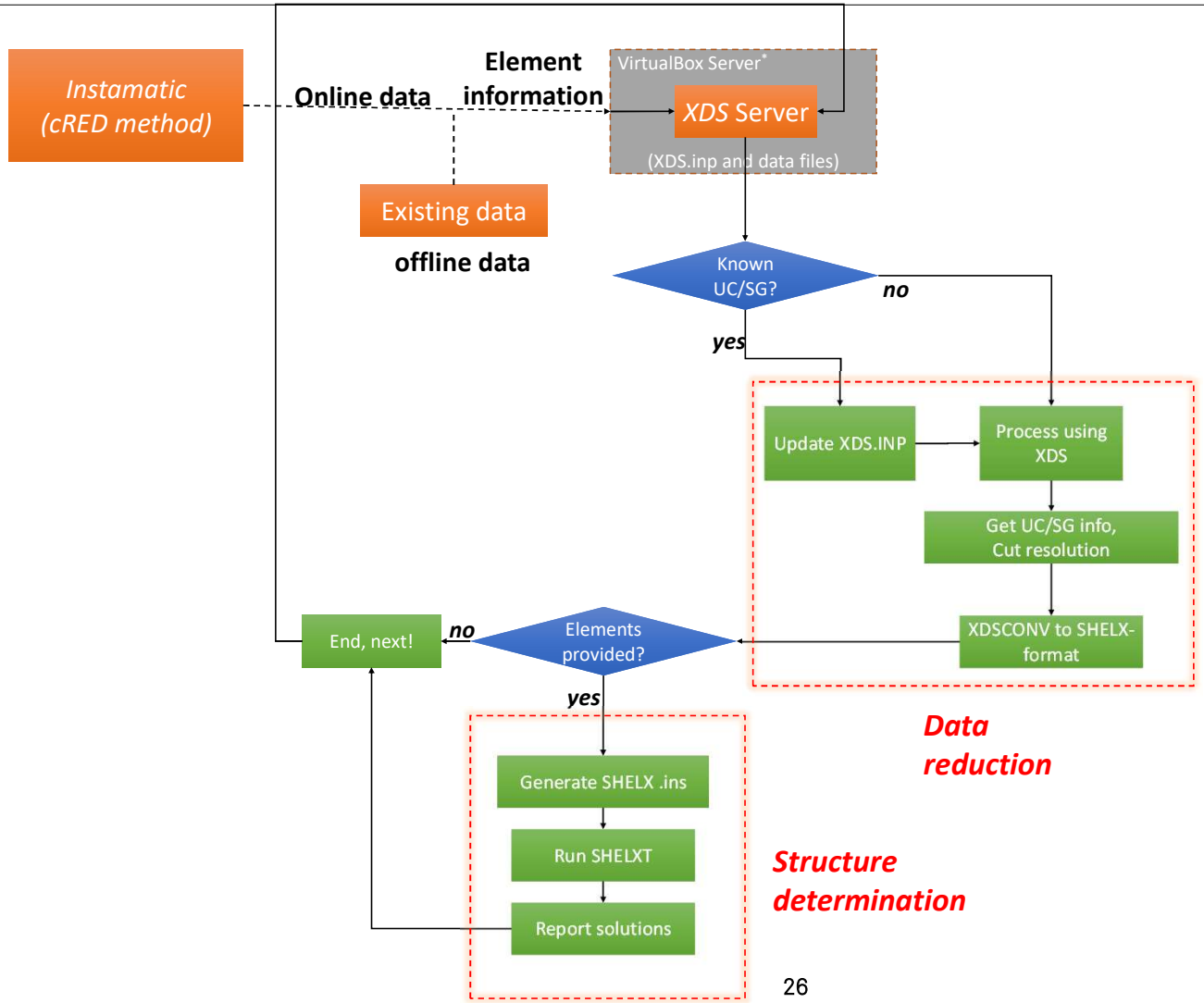








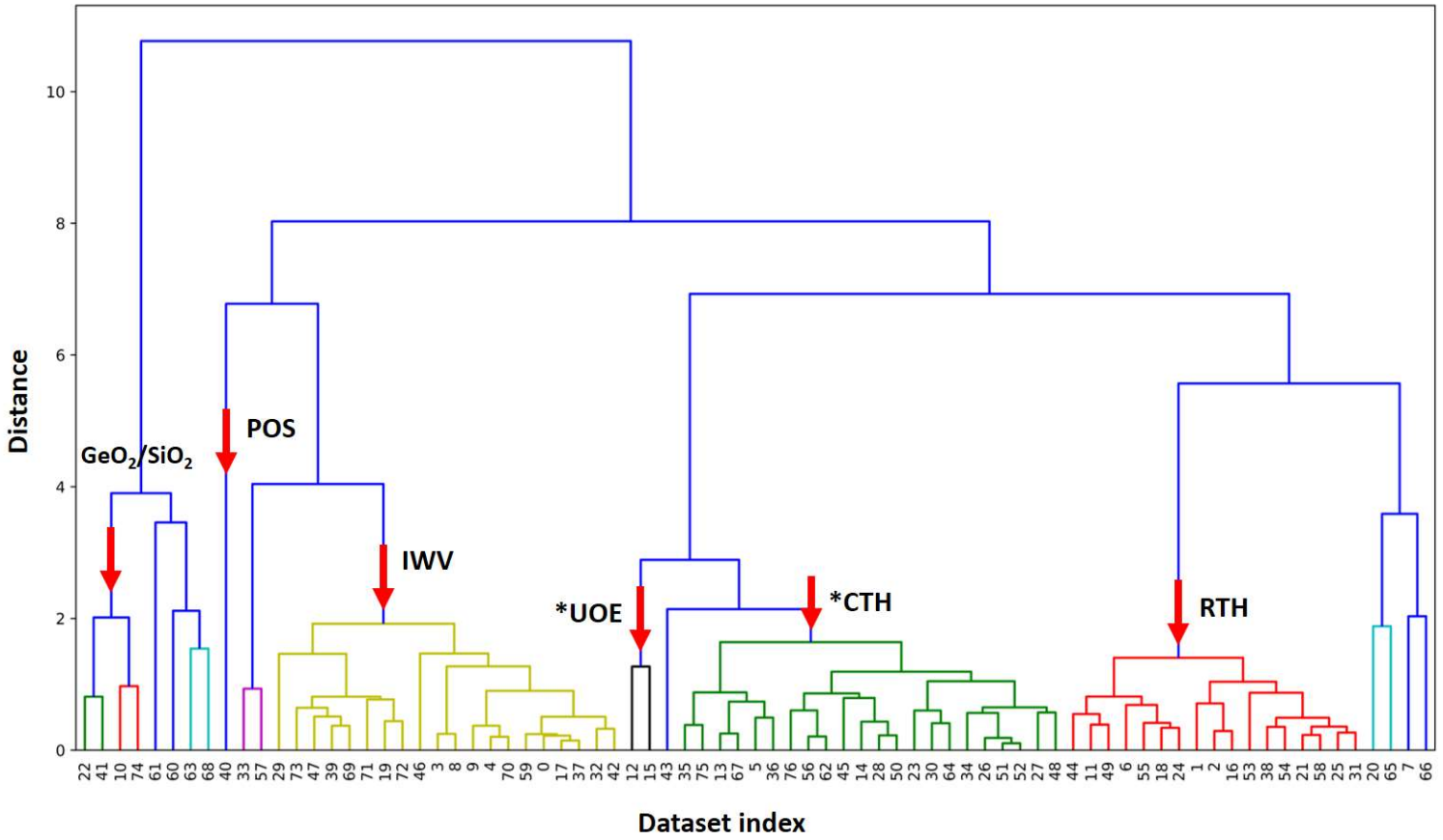
Automated online structure determination pipeline in *Instamatic*



Dr. Bin Wang
Dr. Yi Luo
Prof. Xiaodong Zou

Stockholm
university

Identifying phases in a complex zeolite synthesis (single batch)



Data collection
JEOL 2100 with Timepix
7 hours measurement
-> 321 crystals = ~50/hr

Data
321 data sets
122 > 20° rotation
74 Indexed (DIALS)

Cluster analysis
6 different phases

Yi Luo (Stockholm University)