

Lecture 1. Machine Learning in Full Electron Microscopy Workflow

Sergei V. Kalinin

University of Tennessee, Knoxville and
Pacific Northwest National Laboratory



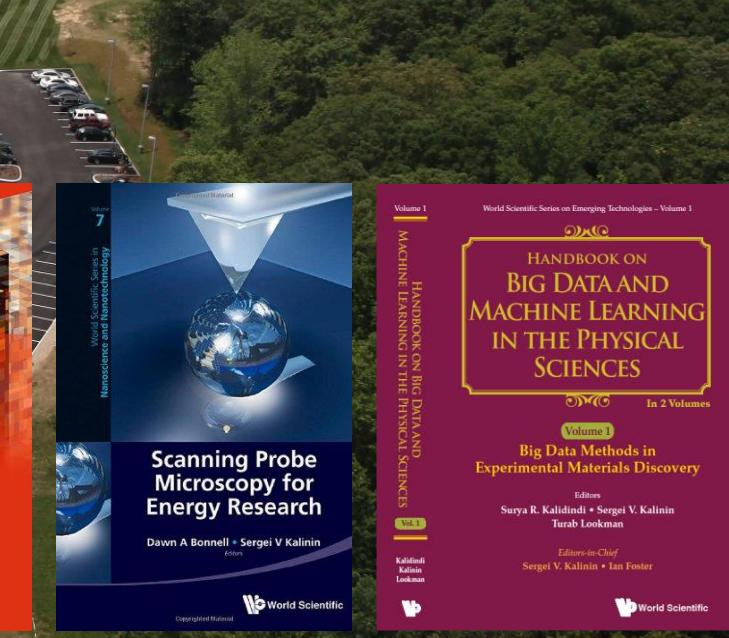
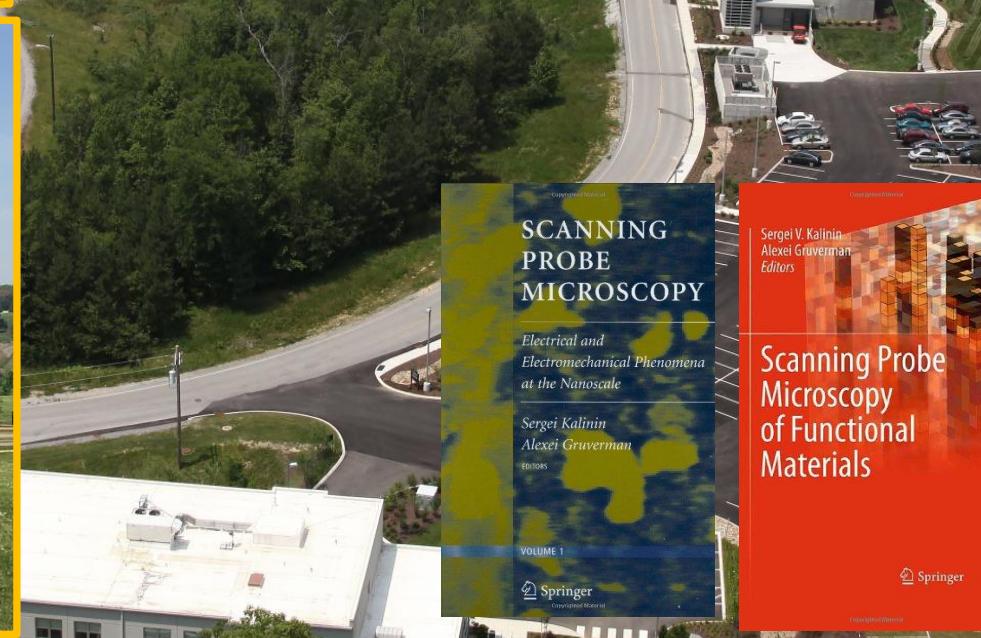
2002 - 2022

Since 2022



2022 - 2023

amazon





Kevin Roccapriore

Founder, AtomQ

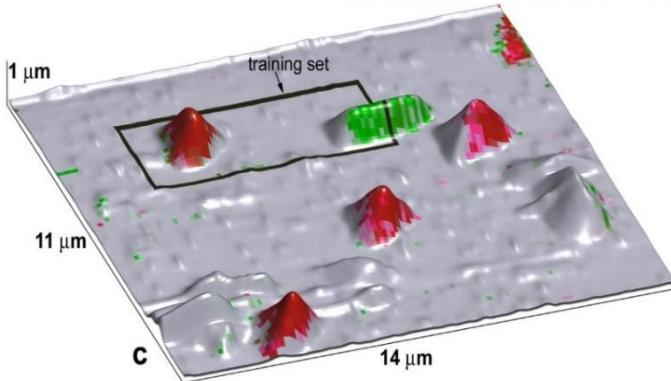
*Build quantum world
one atom at a time*

- Staff scientist at Oak Ridge National Laboratory's Center for Nanophase Materials Sciences (CNMS).
- PhD in Physics from the University of North Texas in 2018, which was nanofabrication for nanophotonics applications.
- Post-doc at ORNL (2018-2021), focusing on machine learning in the scanning transmission electron microscope (STEM) toward automated experiments, including STEM-EELS and 4D-STEM.
- Currently, focuses on ML on-the-fly in the STEM for precise atomic fabrication by site-specific atom targeting with the electron beam, automated experiments learning and utilizing structure-property relationships, and automated multidimensional data acquisition and analysis.

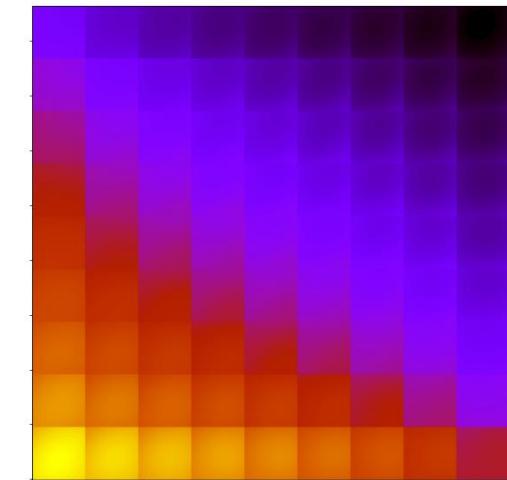
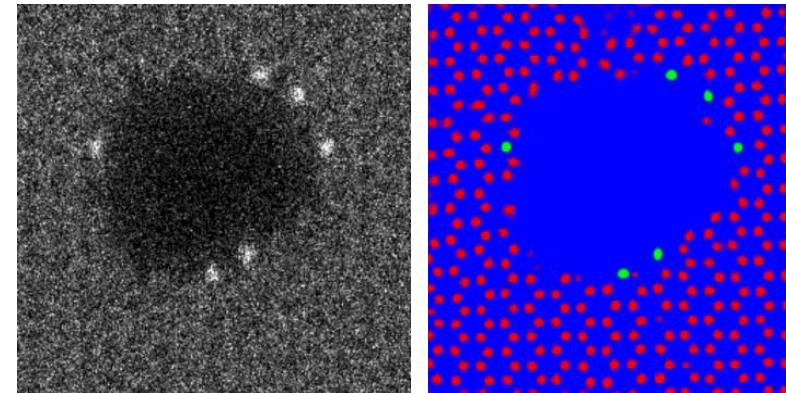
Machine Learning in Microscopy

Deep networks for learning physics

Shallow networks for data analysis



Deep networks for data analysis



2006

2008

2010

2012

2014

2016

2018

2020

2022

Principal component analysis

Neural network recognition

NN-based theory-experiment marching

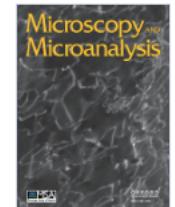
Physics-based ML

Deep learning

- ELIT ML
- Invariant autoencoders

M.P. NIKIFOROV, A.A. VERTEGEL, V.V. REUKOV, G.L. THOMPSON, S.V. KALININ, and S. JESSE, *Functional recognition imaging using artificial neural networks: Applications to rapid cellular identification by broadband electromechanical response*, Nanotechnology **20**, 405708 (2009).

O.S. OVCHINNIKOV, S. JESSE, P. BINTACCHIT, S. TROLIER-MCKINSTRY, and S.V. KALININ, *Disorder identification in hysteresis data: recognition analysis of random-bond random-field Ising model*, Phys. Rev. Lett. **103**, 157203 (2009).



Volume 16, Issue S2

1 July 2010

< Previous Next >

JOURNAL ARTICLE

Towards the Thinking Microscope

[Get access ▾](#)

OS Ovchinnikov, S Jesse, SV Kalinin, HJ Chang, SJ Pennycook, AY Borisevich

Microscopy and Microanalysis, Volume 16, Issue S2, 1 July 2010, Pages 160–161,
<https://doi.org/10.1017/S1431927610062720>

Published 01 August 2010

[Cite](#)[Permissions](#)[Share ▾](#)

Abstract

Extended abstract of a paper presented at Microscopy and Microanalysis 2010 in Portland, Oregon, USA, August 1 – August 5, 2010.

Issue Section: Instrumentation and Techniques Symposia > Aberration-Corrected Electron Microscopy: Exploring Materials Through New Eyes

[Cite](#) [Share](#) [Jump to](#) [Expand](#)

ACS Nano > Vol 15/Issue 8 > Article

REVIEW | July 16, 2021

Automated and Autonomous Experiments in Electron and Scanning Probe Microscopy

Sergei V. Kalinin*, Maxim Ziatdinov, Jacob Hinkle, Stephen Jesse, Ayana Ghosh, Kyle P. Kelley, Andrew R. Lupini, Bobby G. Sumpter, and Rama K. Vasudevan

Big-deep-smart data in imaging for guiding materials design

Sergei V. Kalinin [✉](#), Bobby G. Sumpter & Richard K. Archibald[Nature Materials](#) 14, 973–980 (2015) [Cite this article](#)11k Accesses | 231 Citations | 20 Altmetric | [Metrics](#)

Abstract

Harnessing big data, deep data, and smart data from state-of-the-art imaging might accelerate the design and realization of advanced functional materials. Here we discuss new opportunities in materials design enabled by the availability of big data in imaging and data analytics approaches, including their limitations, in material systems of practical interest. We specifically focus on how these tools might help realize new discoveries in a timely manner. Such methodologies are particularly appropriate to explore in light of continued improvements in atomistic imaging, modelling and data analytics methods.

Machine learning for automated experimentation in scanning transmission electron microscopy

Sergei V. Kalinin [✉](#), Debangshu Mukherjee [✉](#), Kevin Roccapriore, Benjamin J. Blaiszik, Ayana Ghosh, Maxim A. Ziatdinov, Anees Al-Najjar, Christina Doty, Sarah Akers, Nageswara S. Rao, Joshua C. Agar & Steven R. Spurgeon[npj Computational Materials](#) 9, Article number: 227 (2023) [Cite this article](#)

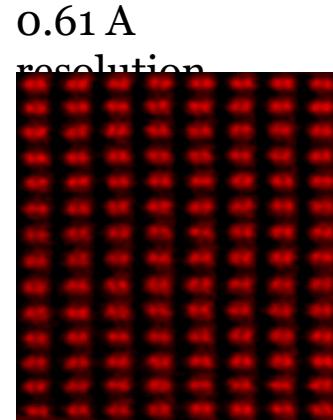
Image atomic columns
Spectra from single atoms



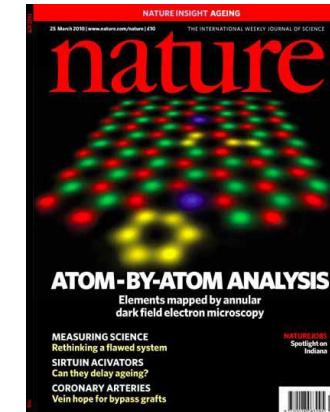
First AC (Nion)

1997

Image light atoms and sensitive materials



Diffraction from subatomic volumes



Beams with orbital momentum



Beam manipulation

Vortex beams

High-res EELS

In-situ holders

4D STEM

Physics extraction

Atomic manipulation

Sample
Source
Detector
Data
Feedback

2002
Prototype correctors

2006
TEAM project

Broad adoption of AC STEM

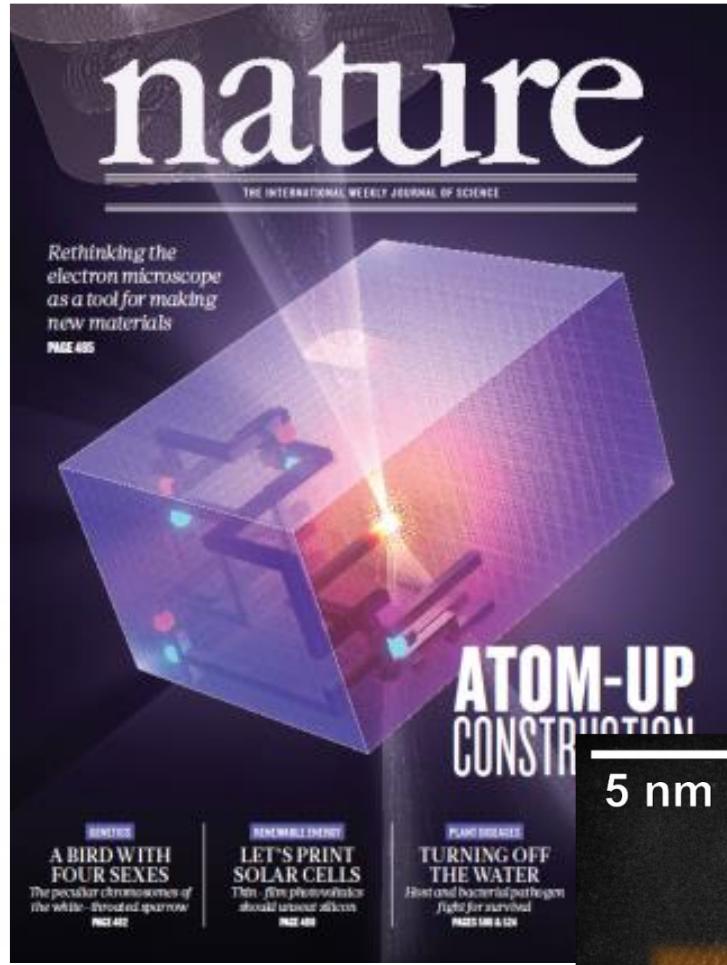
2012
Segmented 4D STEM detectors

2016
Vortex beams

2018

Images from A. Lupini

Electron Microscopy yields large volumes of spatially-resolved data on atomic coordinates (**bond length** and **bond angles**), low energy and core-level **excitations**, and scattering from subatomic volumes. In addition, electron beams can be used to **manipulate matter atom by atom**.



Volume 75, Issue 6
June 2022



< Previous Article Next Article >

A quantum lab in a beam FREE

Advances in electron microscopy have revolutionized atomic-scale imaging, characterization, and manipulation of materials.

Sergei V. Kalinin; Stephen Jesse; Andrew R. Lupini



Physics Today 75 (6), 30–36 (2022);

<https://doi.org/10.1063/PT.3.5018>



PDF



Share



Reprints and Permissions

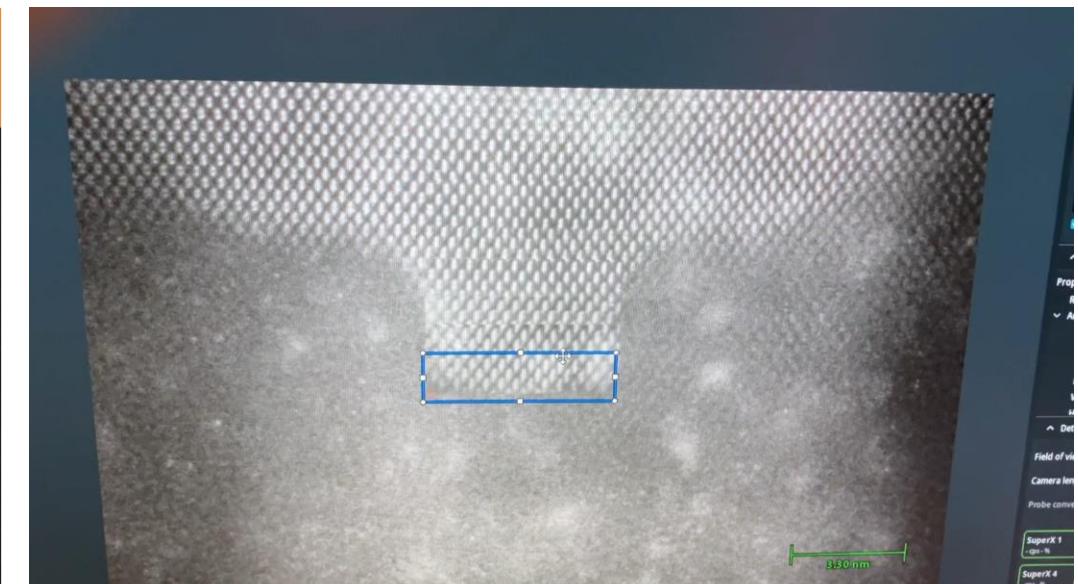
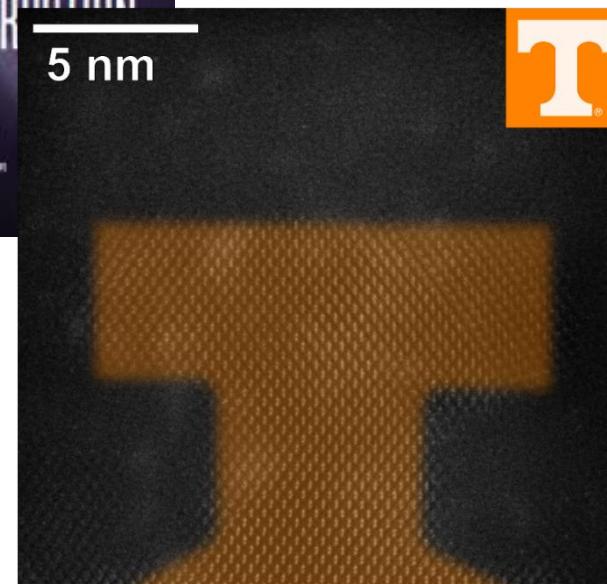


Cite

Topics

[Electron spectroscopy](#), [Scanning electron microscopy](#), [Scanning tunneling microscopy](#), [Transmission electron microscopy](#), [Quantum structures](#), [Quantum mechanical systems and processes](#)

The history of science is filled with questions about the nature of matter, its constituent elements, how properties emerge from the elements' arrangements, and how the arrangements can guide or be guided by energy flows. The answers to those questions have progressed from philosophical proposals of atomic theory to practical demonstrations of atoms' existence to modern quantum theory. And, importantly, the answers have been based on experimental measurements.



ML for Automated Microscopy?

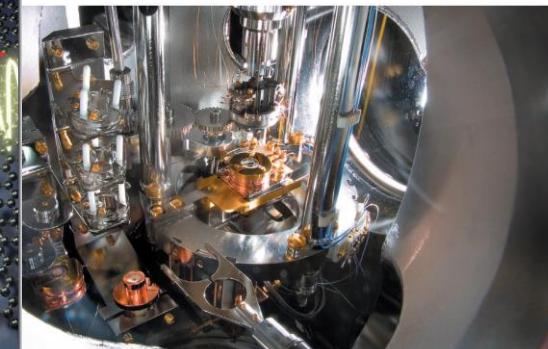
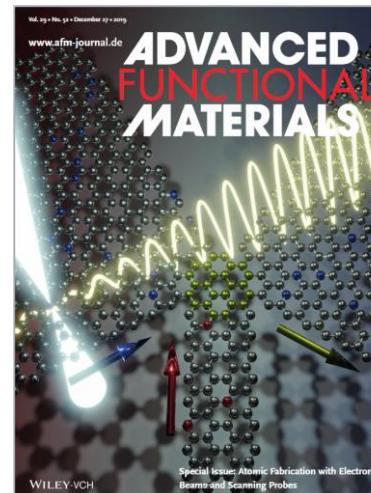
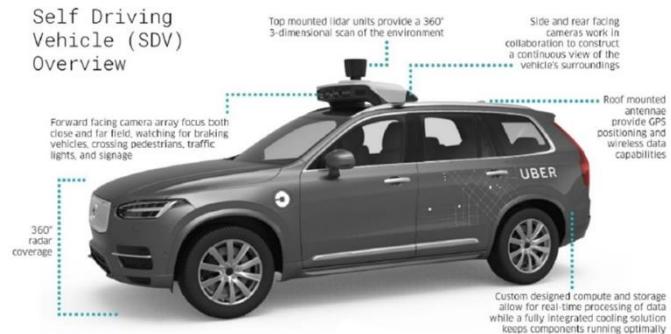
Microscopy today:

- Primary component of research in materials, physics and biology
- 1000s of high-end (S)TEM platforms, ~10,000 overall
- 1000s of high-end UHV SPMs, >50,000 ambient
- Chemical and mass-spectrometric imaging

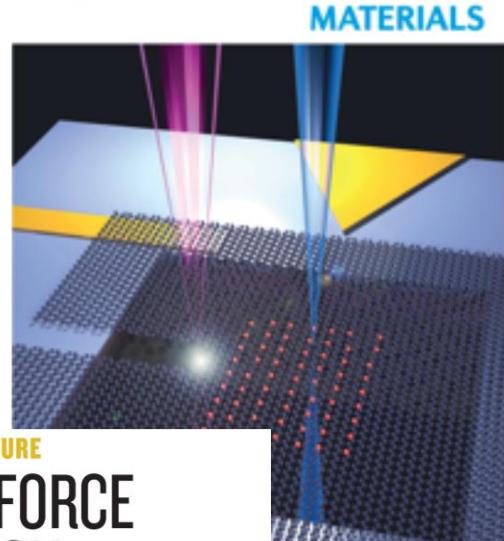
What do microscopists do?

- Most of the time - sit alone in the dark room and turn knobs 😊
- Limited amount of collected data
- Case for automation: CryoEM

Unsurprisingly, inspired by autonomous cars, etc. – multiple proposals to make automated microscopes!



nature
REVIEWS
July 2019 volume 8 no. 7
www.nature.com/nature-reviews



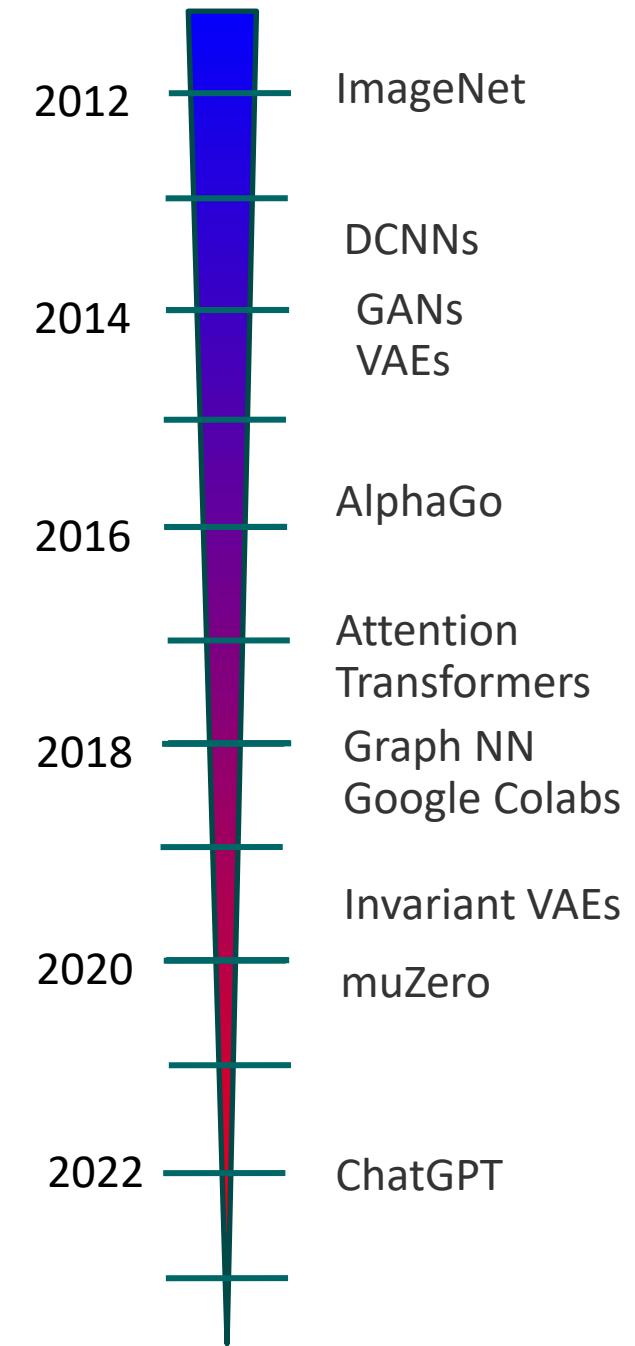
July 2019

Why Machine Learning?

- Last decade has experienced an explosive growth of machine learning and artificial intelligence applications
- These developments have spanned areas from computer vision to medicine to autonomous systems and games
- However, the progress and impact as applied to experimental physical sciences has been minimal....

Why is it difficult?

- Requires domain expertise and domain-specific goals
- Deeply causal and hypothesis drive nature of domain sciences
- No single answer: culture, not a method
- Infrastructure, open code, open data
- **Most important:** active nature of scientific process



scikit-learn

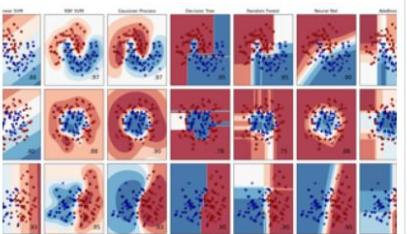
Machine Learning in Python

Getting Started Release Highlights for 1.5

Classification

Identifying which category an object belongs to.

Applications: Spam detection, image recognition.
Algorithms: Gradient boosting, nearest neighbors, random forest, logistic regression, and more...

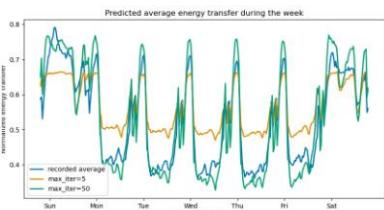


Examples

Regression

Predicting a continuous-valued attribute associated with an object.

Applications: Drug response, stock prices.
Algorithms: Gradient boosting, nearest neighbors, random forest, ridge, and more...

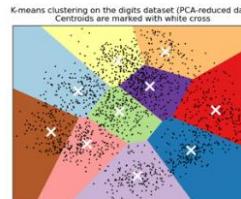


Examples

Clustering

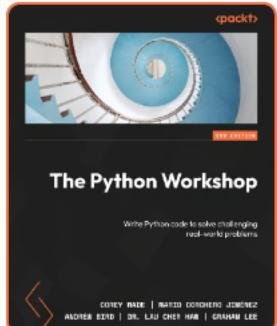
Automatic grouping of similar objects into sets.

Applications: Customer segmentation, grouping experiment outcomes.
Algorithms: k-Means, HDBSCAN, hierarchical clustering, and more...



Examples

Home > Programming > Programming Language > The Python Workshop Second Edition - Second Edition



The Python Workshop Second Edition: Write Python code to solve challenging real-world problems, Second Edition

By Dr. Lau Cher Han , Andrew Bird , Mr. Mario Corchero Jiménez , Mr. Corey Wade , Graham Lee

\$41.99 \$28.99

★★★★★ 4.7 (3 Ratings)

Book • Nov 2022 • 600 pages • 2nd Edition

eBook
\$41.99
\$28.99

Print
\$51.99

Subscription
Free Trial
Renews at \$15.99/p/m

- Packtpub.com
- Manning.com
- GitHub

AtomAI latest

Search docs

NOTES

ReadMe
LICENSE
Scientific output

PACKAGE CONTENT

AtomAI Models
Trainers and Predictors
Neural Nets
Loss Functions and Metrics
Other utilities

EXAMPLES

Colab notebooks

ON-DEMAND WEBINAR How to reduce your organization's reliance on "bad" open source packages [WATCH NOW](#) TIDELIFT

On-demand webinar: Reduce your organization's reliance on "bad" open source packages. [WATCH NOW!](#)

Ad by EthicalAds • i

/ Welcome to AtomAI's documentation!

Welcome to AtomAI's documentation!

Notes

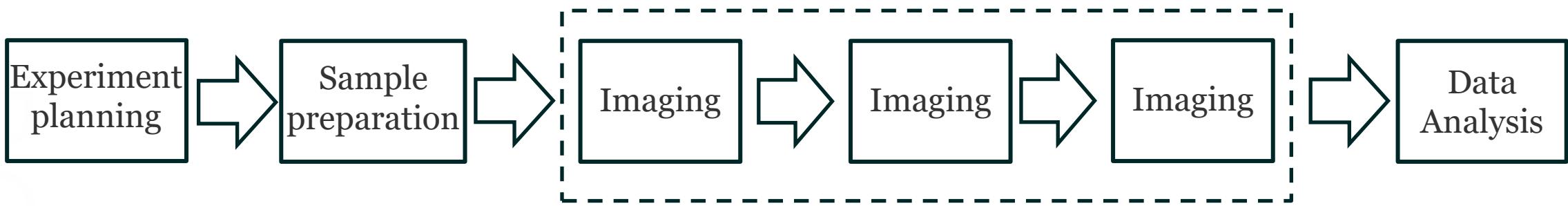
- ReadMe
- What is AtomAI
 - Why AtomAI
- How to use it
 - Semantic segmentation
 - ImSpec models
 - Deep ensembles
 - Variational autoencoders (VAE)
 - Custom models
 - Not just deep learning
- Installation
- LICENSE
- Scientific output

Package Content

- AtomAI Models
 - Segmentor
 - ImSpec
 - Variational Autoencoder (VAE)
 - Rotational Variational Autoencoder (rVAE)
 - Joint Variational Autoencoder (jVAE)
 - Joint Rotational Variational Autoencoder (jrVAE)
 - Deep Kernel Learning
 - Load trained models

What is your goal?

Level 5: Downstream Use of Microscopy Data:
Incorporates microscopy data into theory analysis pipelines, closing the synthesis-characterization-discovery loop.



Level 4: Upstream Task Planning:
ML is used for planning experiments, including sample selection and integrating microscopy with materials synthesis.

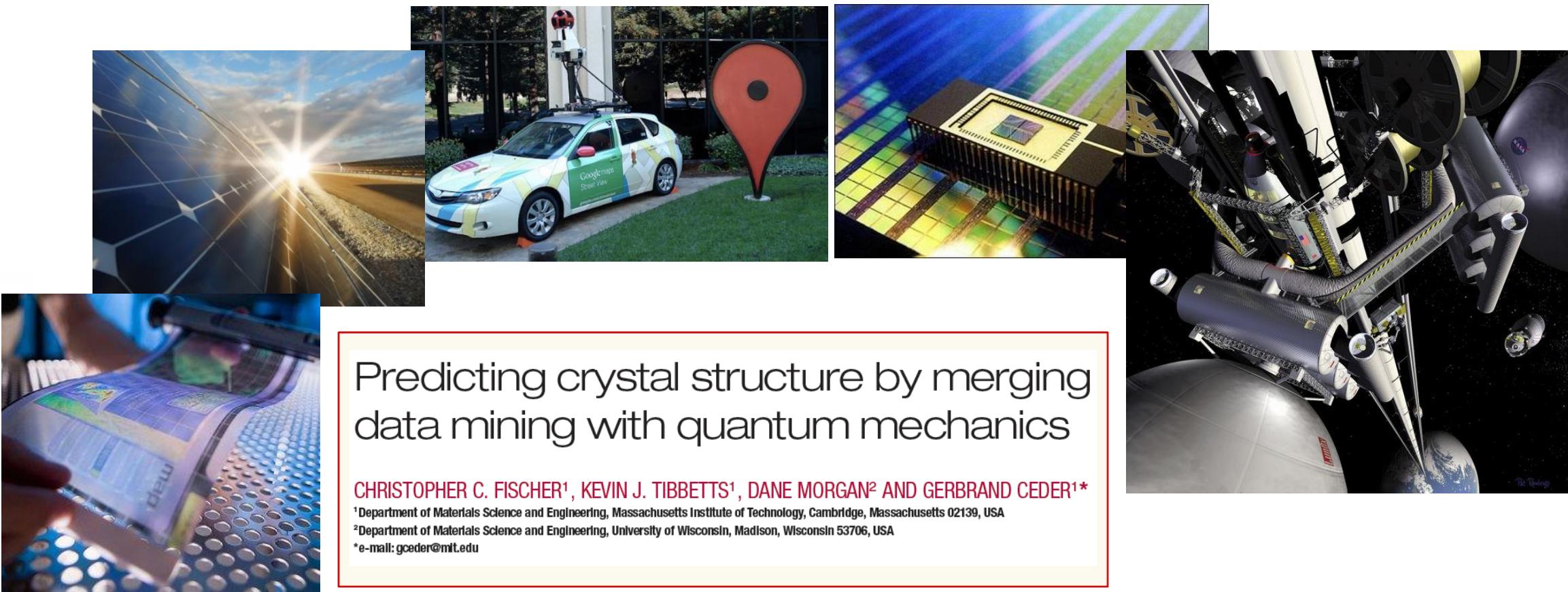
Level 2: Real-Time Analytics

- ML helps represent data in a form that is more understandable to humans.
- Decisions are still made and orchestrated by humans.

Level 1: Post-Acquisition Data Analysis

Level 3: ML Agent Introducing Decisions: automated microscopy

The World is a Material Opportunity



Predicting crystal structure by merging
data mining with quantum mechanics

CHRISTOPHER C. FISCHER¹, KEVIN J. TIBBETTS¹, DANE MORGAN² AND GERBRAND CEDER^{1*}

¹Department of Materials Science and Engineering, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA

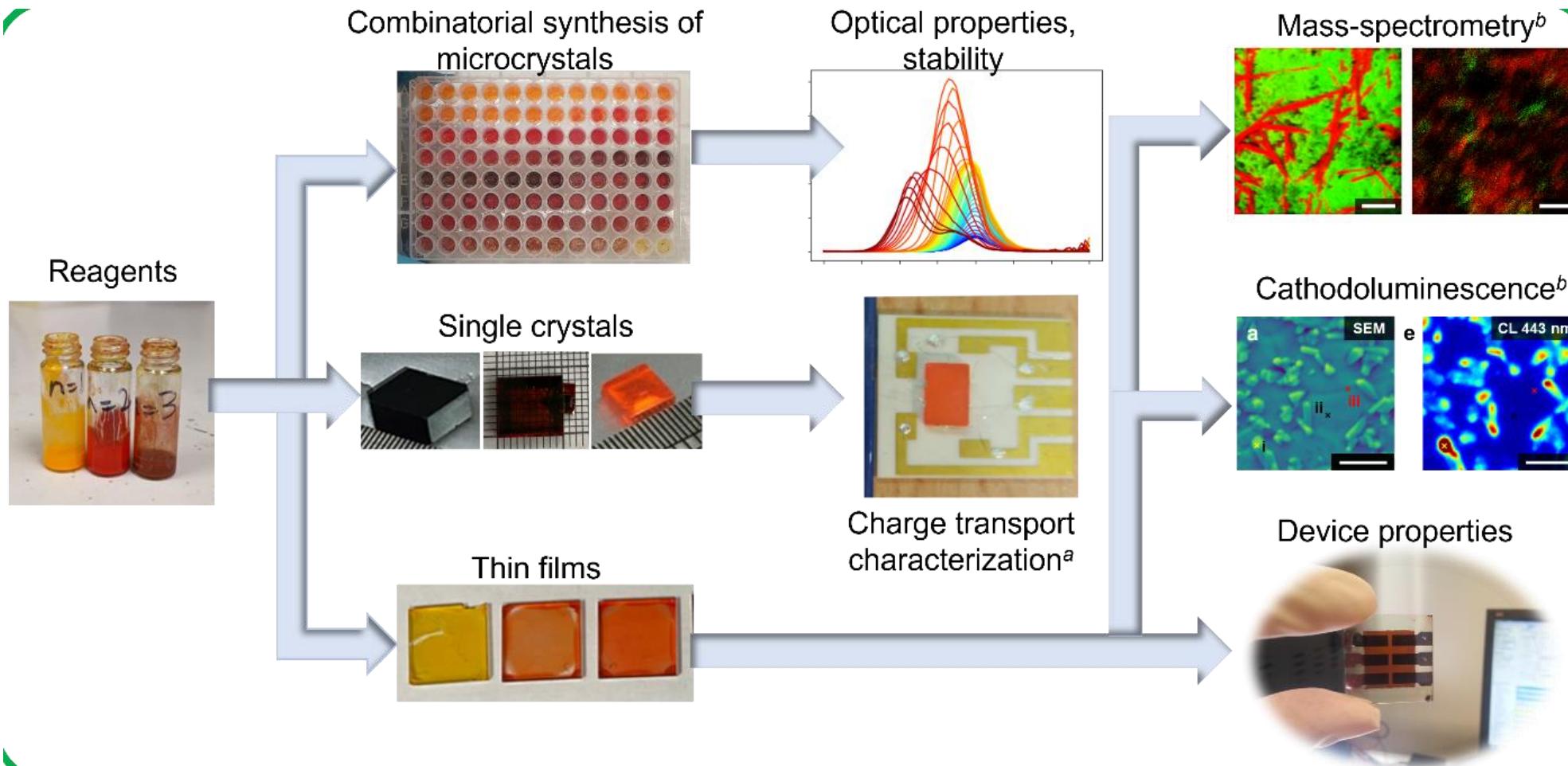
²Department of Materials Science and Engineering, University of Wisconsin, Madison, Wisconsin 53706, USA

*e-mail: gceder@mit.edu

- “**Improve**”: Renewable energy, structural materials
- “**Discover**”: RT superconductivity, high mechanical stress materials
- “**Engineer**”: Quantum computing, single-atom catalysts, biomolecules

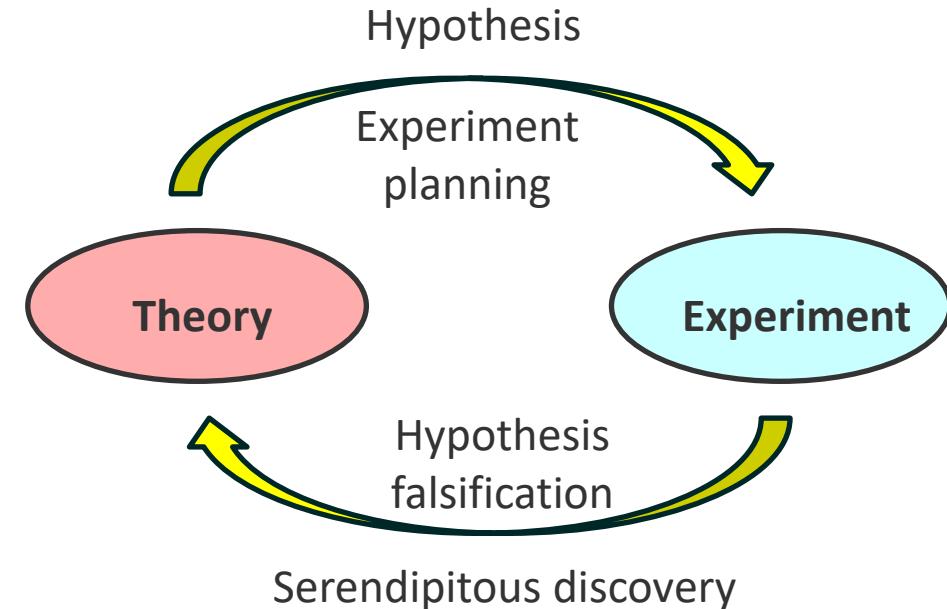
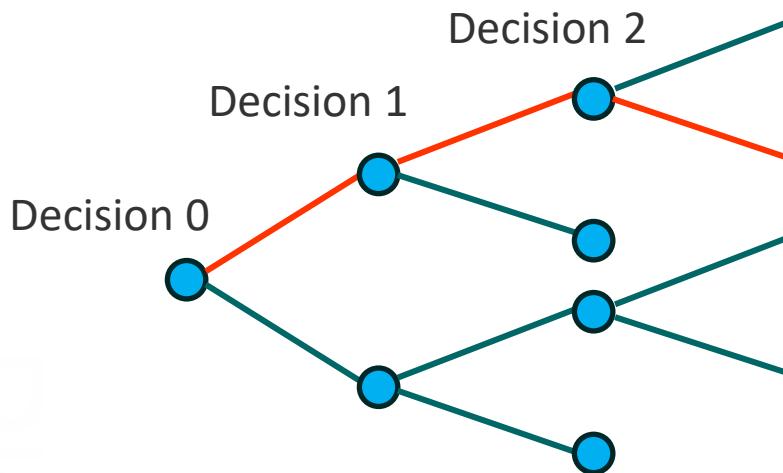
Functionality, manufacturability, cost

Decision science of workflows



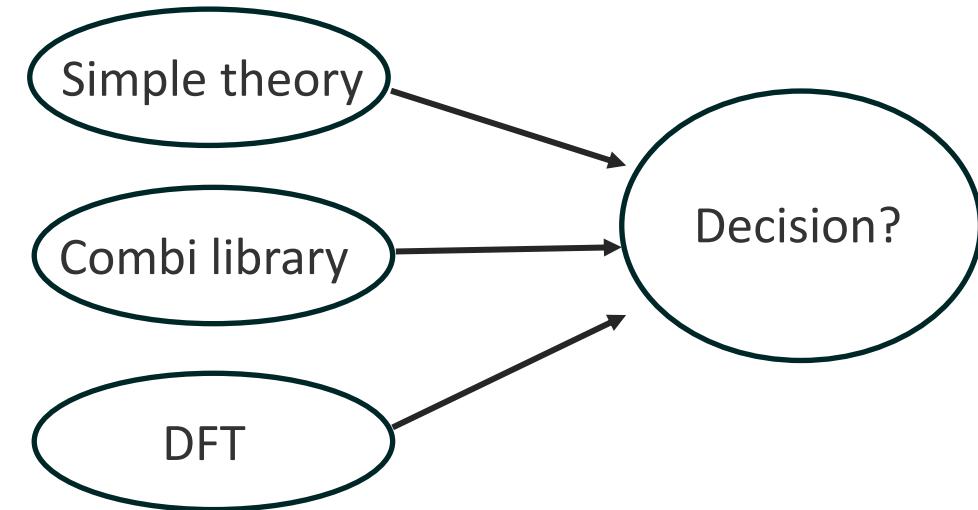
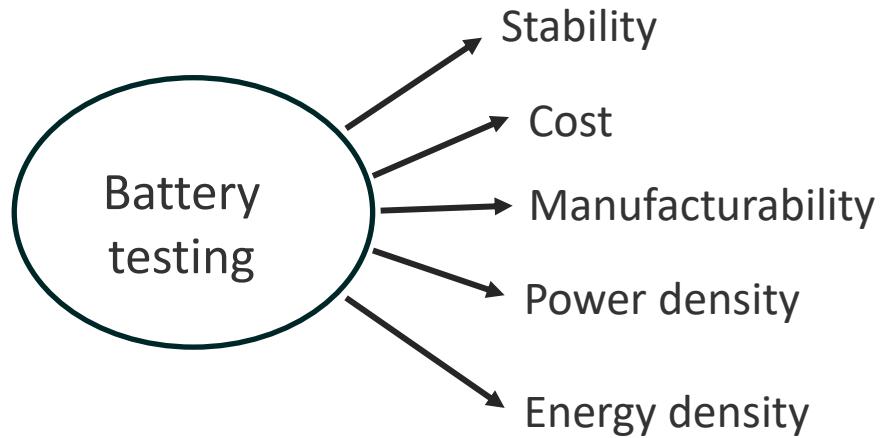
- Multiple levels of decision making based on **perceived gain**, **latencies**, and **costs**
- Iterative cycles between low-cost and expensive measurements
- Learning **basic science/models** as a strategy to minimize cost and answer interventional and counterfactual questions

Decision science of workflows



- **Experiment is a combinatorial space of opportunities:**
 - Investing only in scaling of throughput is only a linear improvement
 - **Knowledge of physics often allows to reduce complexity: combinatorial to linear:**
 - Basic science pays off (with time)!
 - **Science is a cycle between theory-driven hypothesis generation and experiment:**
 - We need to accelerate all elements of the cycle
 - **Experimental and computational tool development:**
 - Currently constrained by human paradigm
- If the part of a workflow is automated, our autonomous decision-making ability should match the level of autonomy!

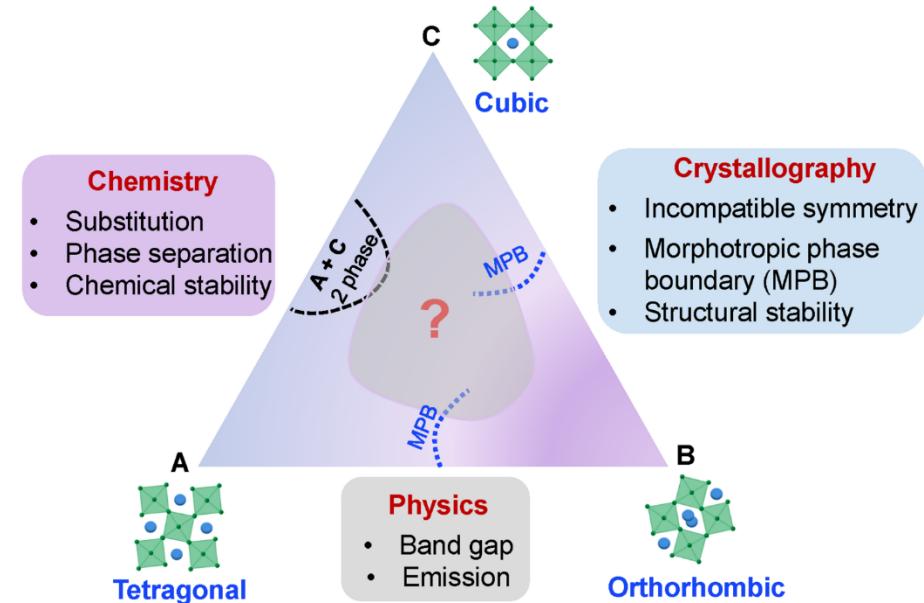
Decision science of workflows



1. We need to balance multiple functionalities
2. Integrate multiple sources of data
3. Make decisions considering costs, latencies, physical inferential biases, and beliefs

Key consideration: reward function

1. Pure physical discovery (symbolic laws)
2. Data-driven exploration
3. Materials optimization
4. ...

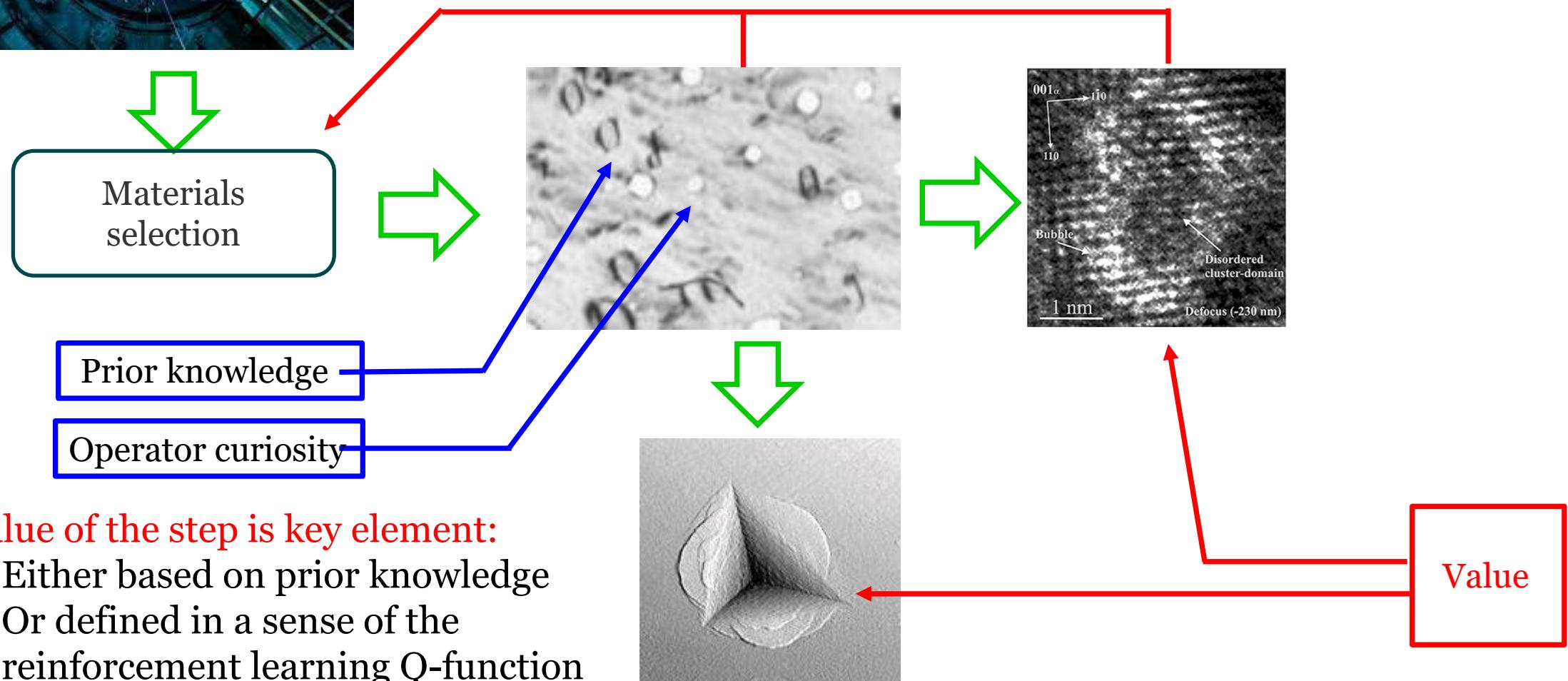


Workflows for Nuclear Materials Design



Traditional experiment:

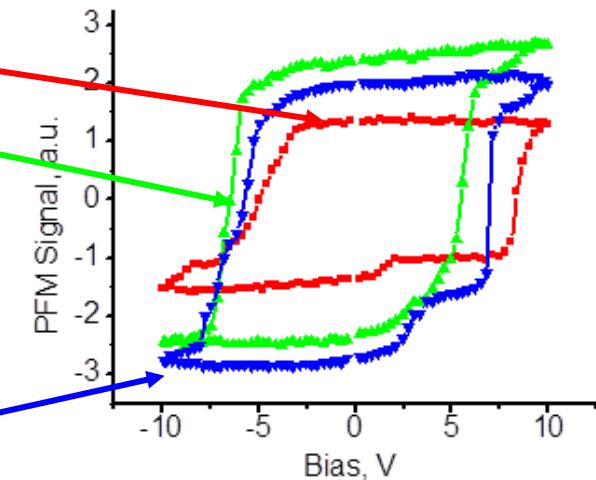
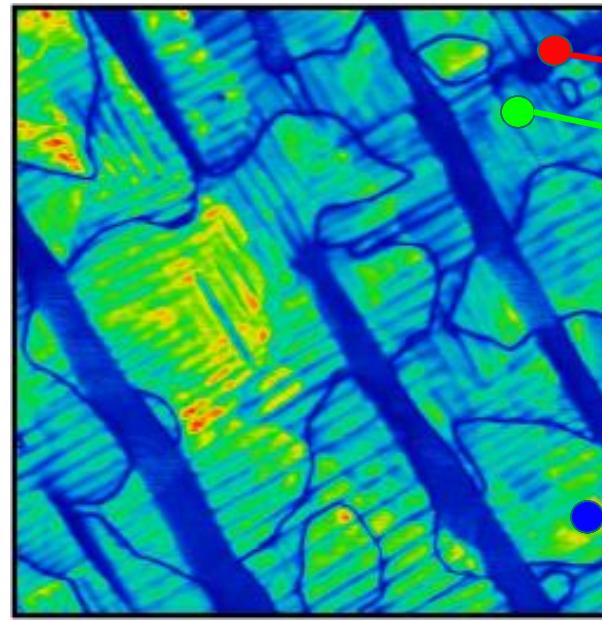
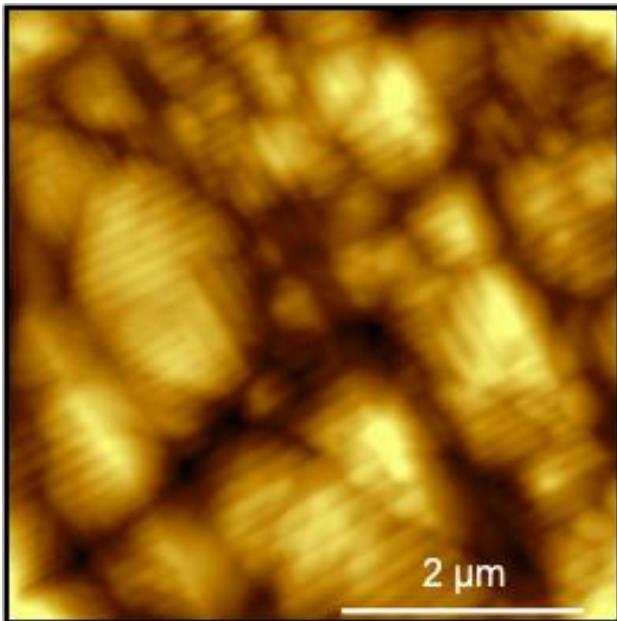
1. Always based on workflows
2. Ideated, orchestrated, and implemented by humans
3. The “gain of value” during the workflow implementation is uncertain



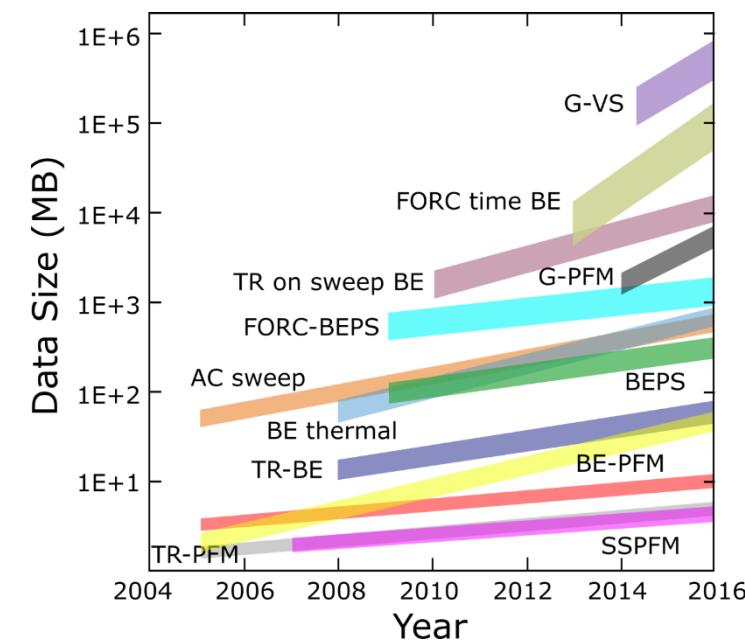
Value of the step is key element:

- Either based on prior knowledge
- Or defined in a sense of the reinforcement learning Q-function

Decision Making in SPM



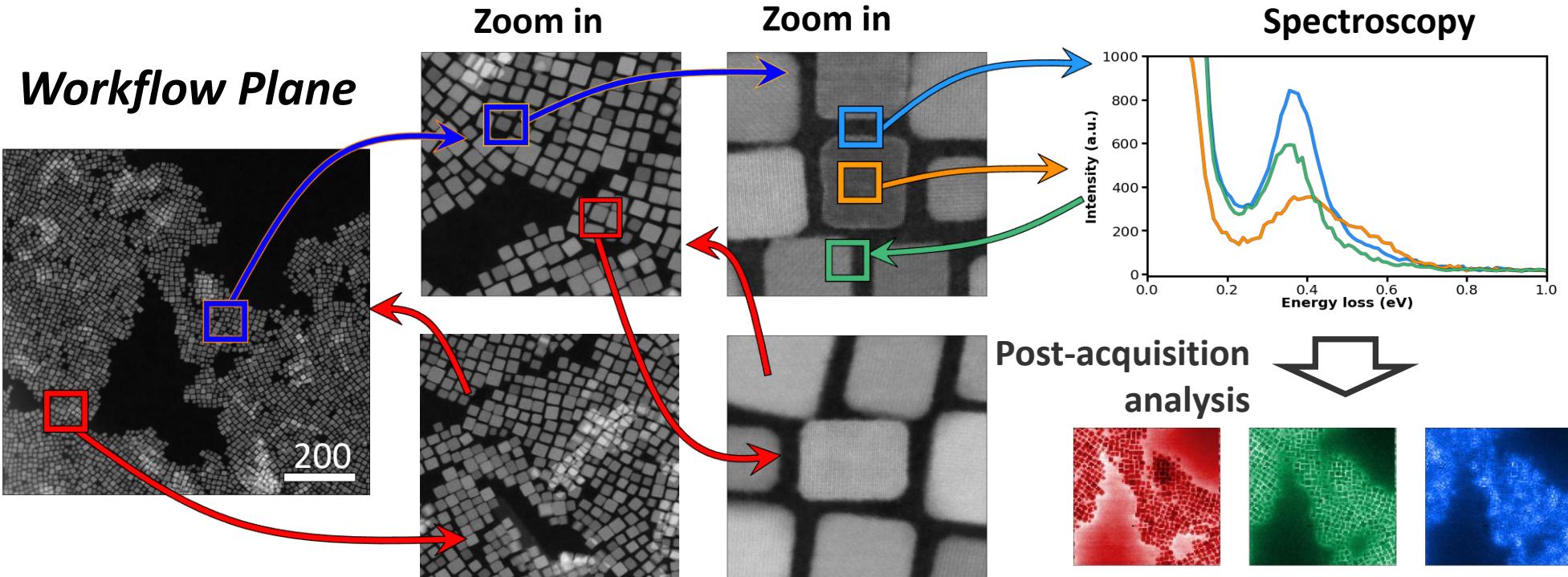
- Interesting functionalities are expected at the certain elements of domain structure
- We can guess some; we have to discover others
- **Experimental objectives → ML Rewards**
 - Microscope optimization
 - Properties of a priori known regions of interest
 - Discovery of regions with interesting properties
 - Physical theory falsification



Objective and Reward

Workflows in STEM

Prior Knowledge



Instrument Plane

Minimal instruction set control language

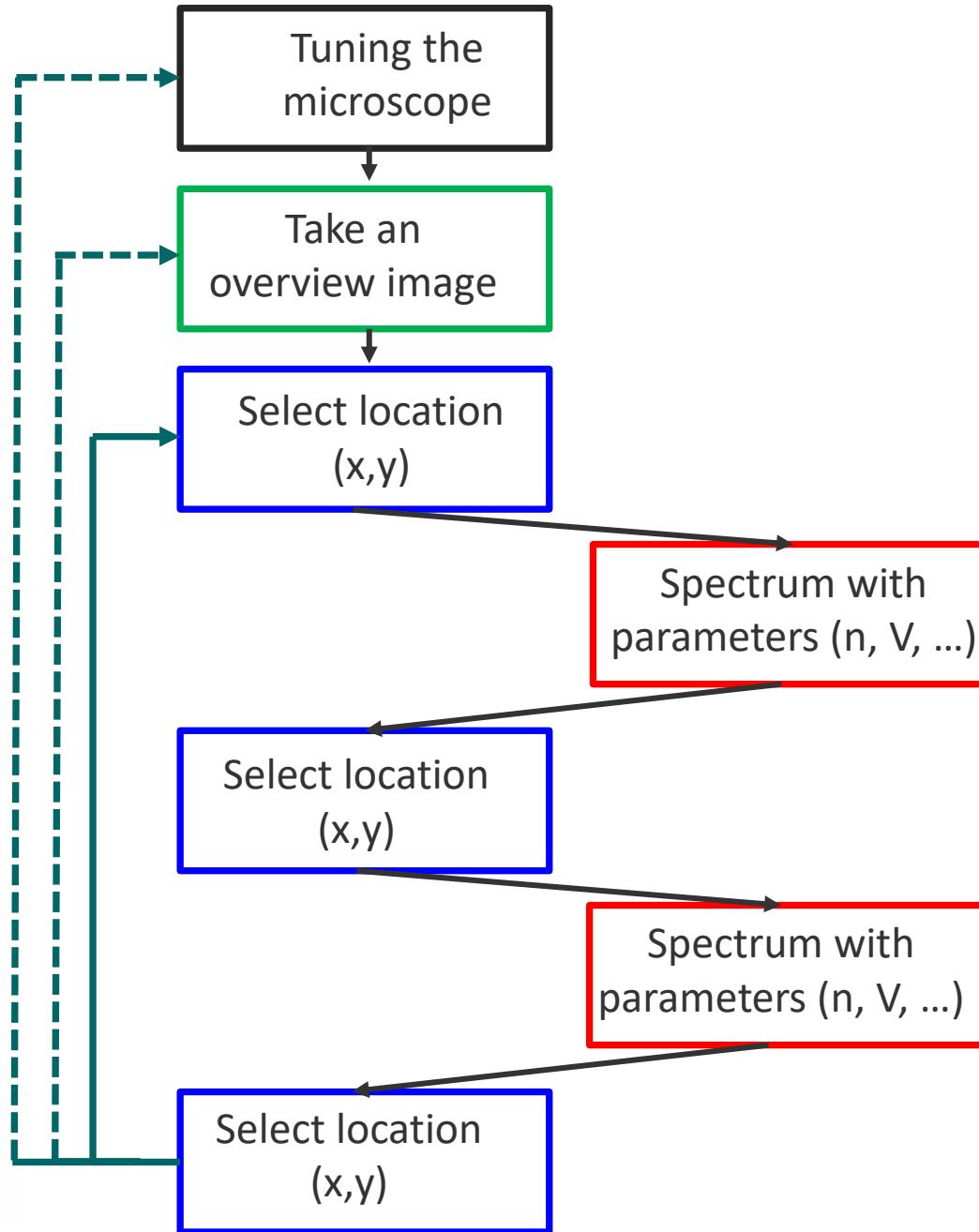
Load sample
and
tune microscope
etc.

Overview scan
and
tune parameters

Initiate scan

Position probe
(x, y)

Initiate spectrum
(x, y, v)



To implement the ML workflows, we start from emulating the human operations:

- Well defined and explainable commands
- Extensive domain expertise
- Potentially available data from experiments

Development of ML workflows can give rise to more complex imaging modalities

- Data volumes and dimensionalities above human level
- More complex modes of sampling
- “Guardian angel” modules

However, we always have to think about

- Reward function(s) for imaging problem
- Reward functions for materials problem
- Overall objective

Challenges for the Automated Experiment

Elements of realistic workflow design

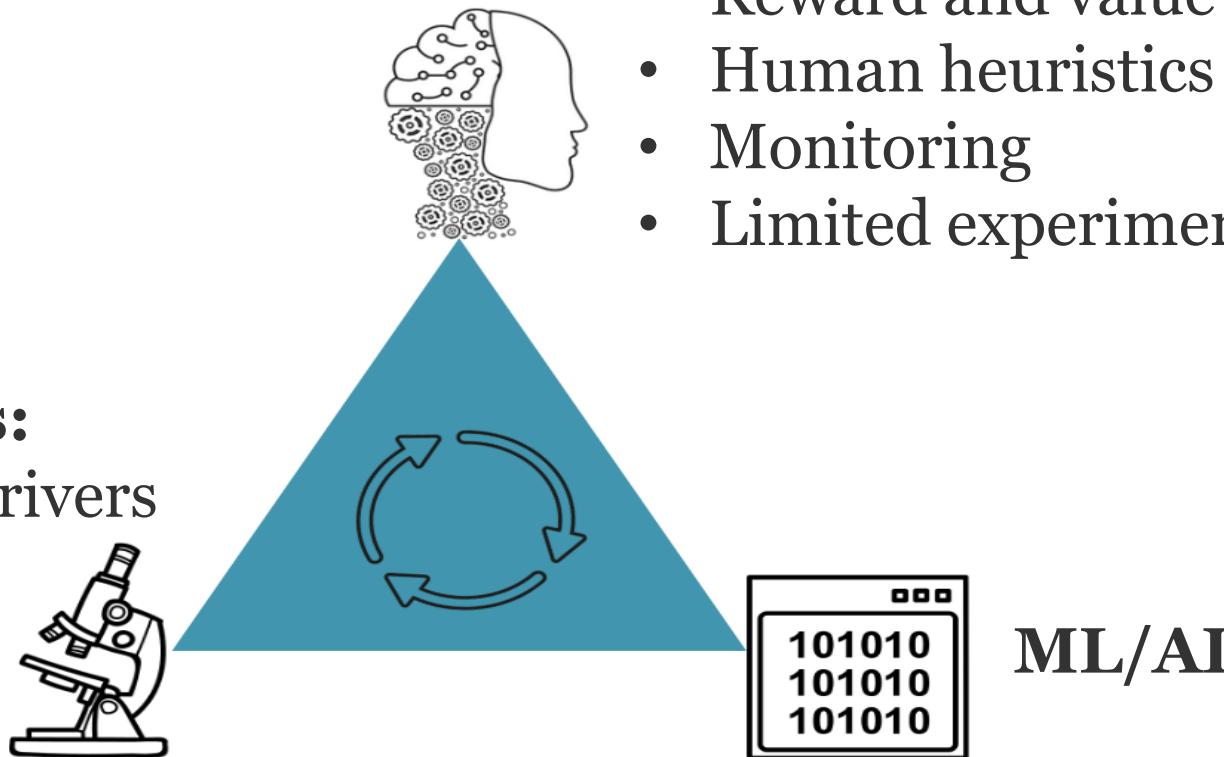
- Co-orchestration of multiple measurement modalities
- Building theory in the loop
- Integration between multiple domains

Workflow design:

- Reward and value functions
- Human heuristics
- Monitoring
- Limited experimental budget

Engineering controls:

- Instrument-specific drivers
- Hyperlanguage
- Python APIs



DigitalMicrograph

File Edit Display Process Analysis Window Microscope Spectrum EELS EFTEM SI Volume Custom Camera Help

Microscope Microscope Display

Microscope System

200 kV
SCANNING, nP, Spot 7
Mag: x 115k CL: 91 mm

Scan Valve Blank

ADF BF B-Stop

EFTEM EELS EF-CCD

Energy Loss: 0.0 eV
Hi-SNR aperture, 0.75 eV/Ch

GATAN

Tune GIF

Find ZLP AutoFocus

Calibration

Camera Monitor

Temperature 5.0 C

Health Status

Stage Tracker

Output

Output Image Browser Script Debugger

Vertical pixel Step

Vertical Spacing

Zoom factor

ds_para [-16384.000244140625, -16384.000244140625, 409.6000061035156]

FilterControl

Main Adjust Calibrate

GIF Continuum ER

Primary Energy 200.0 keV

Shift 0.0 eV

Adjust 0.0 eV

HT Offset 0.0 eV

Slit In Width 900.0 eV

Dispersion 0.75 eV/Ch

Mode Spectroscop

Aperture 5 mm

Drift Tube 0.0 eV

Wobble 0.0 eV

Technique Manager

STEM SI

Scan

Spot Focus Rotate

View Pixel Time (ps): 63.42 Search Prev

STEM Alignment

SI Acquisition

EELS BF-DF

2D Array Multi-Point

Line Scan Time Series

EF-CCD Camera

EELS

B=49.9 mrad Single Dual

HS+ HS HQ User

ZLP-lock Energy (eV): 0.0

View Exposure (s): 0.01 auto

Capture Frames: 10

Elemental Quantification

EELS Analysis

Zero-Loss Thickness Splice Deconvolve

ColorMix

Jupyter ISAAC_smart_eels_using_edge_detect Last Checkpoint: 23 hours ago

File Edit View Run Kernel Settings Help

```
array_list, shape, dtype = array_server.get_eels()
array = np.array(array_list, dtype=dtype).reshape(shape)
plt.figure()
plt.plot(array)
# plt.ylim(0,1e6)
```

detect edges : do eels on those coordinates

```
import numpy as np
import matplotlib.pyplot as plt

def detect_bright_region(image):
    # Calculate the gradient in the X and Y directions
    gx = np.gradient(image, axis=1) # Gradient in X direction
```

ShareX 15.0

Capture > explorer_nFOIVdh... YWvOPrPkus.png Bmp6leJ6nm.mp4

Upload

Workflows

Tools

After capture tasks

After upload tasks

Destinations

Application settings...

Task settings...

Hotkey settings...

Screenshots folder...

History...

Image history...

Debug

Donate...

Twitter...

Discord...

About...

```
array_server.create_camera()
scale = int(2**14/image_size)
line_p = np.zeros([image_size, image_size, array.shape[0]])
```

```
accepted = 0
for i, y in enumerate(range(image_size)):
    print("line scan ", y, )
    for j, x in enumerate(range(image_size)):
        if edges_detected[i,j]> threshold_eels: # condition to do eels
            accepted+=1
            array_server.set_beam_pos(x, y)
            array_server.acquire_camera()
            array_list, shape, dtype = array_server.get_eels()
            array = np.array(array_list, dtype=dtype).reshape(shape)
            #plt.plot(array)
            line_p[i,j] = array # summing eels to get bright field pixel value
tend = time.time()

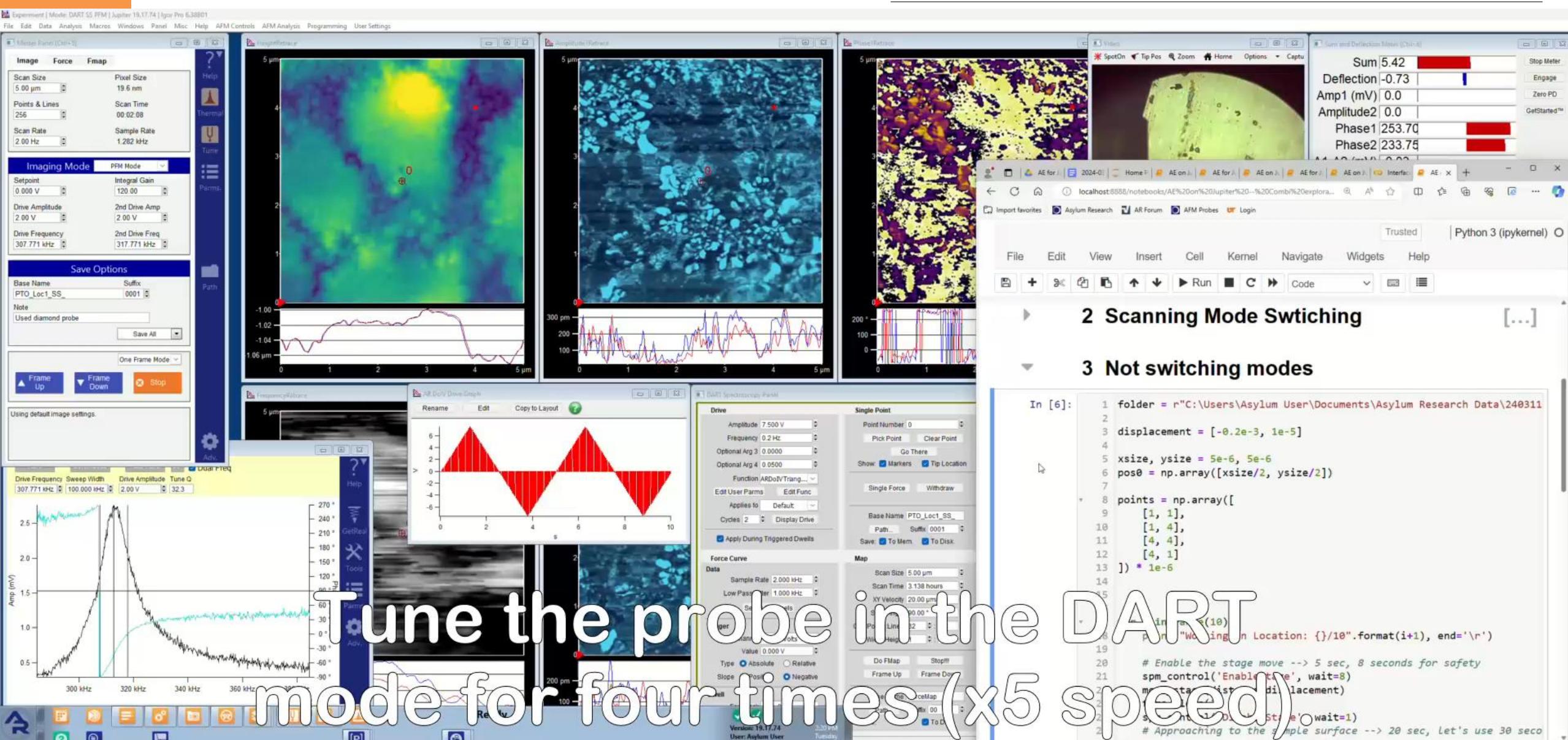
print("accepted_points",accepted)
print("time taken in seconds", tstart - tend)
```

```
# get current position to do eels
#Activate camera
array_server.activate_camera()
array_list, shape, dtype = array_server.get_ds(80)
im_array = np.array(array_list, dtype=dtype).reshape(shape)
plt.figure()
plt.imshow(im_array, cmap="gray")
plt.show()
```

```
fig, (ax1, ax2) = plt.subplots(1,2)
ax1.imshow(line_p.sum(axis=2))
```

freeilm remote xbox_rec Record game game Xbox game My Inter Home ISAAC_

Acquire a digiscan image on Electron Microscope from Supercomputer



Tune the probe in the DART mode for four times (x5 speed)

File Edit View Insert Cell Kernel Widgets Help

27
28
29

30 move_(-volt*2-(offsetvx), 0, 0-offsetvy, 0, move_speed)

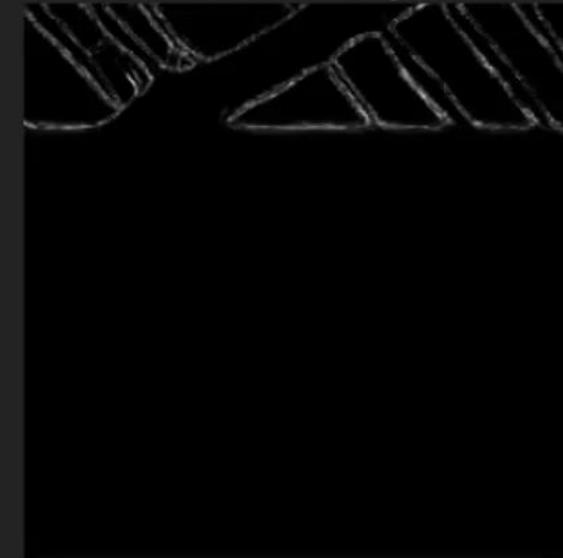
Amplitude



Ferroelastic Walls



Uncertainty



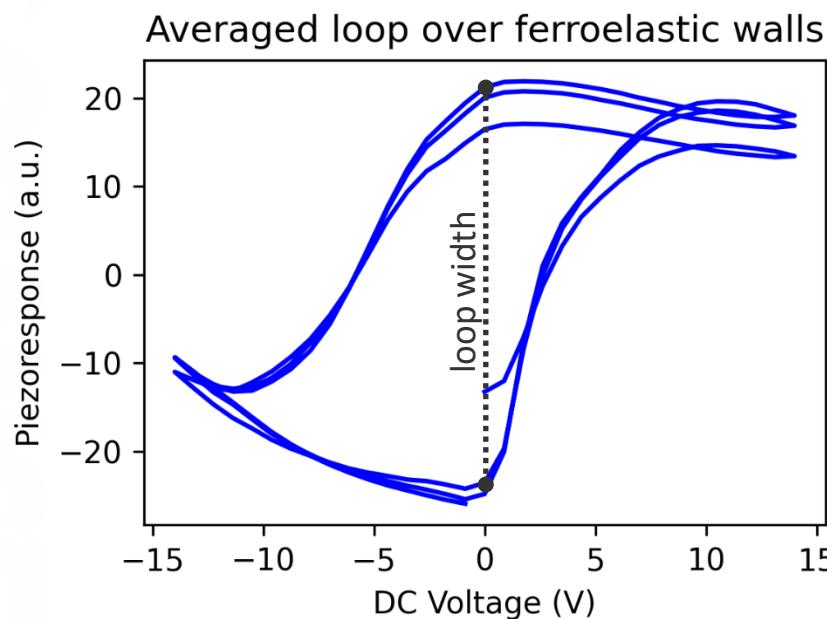
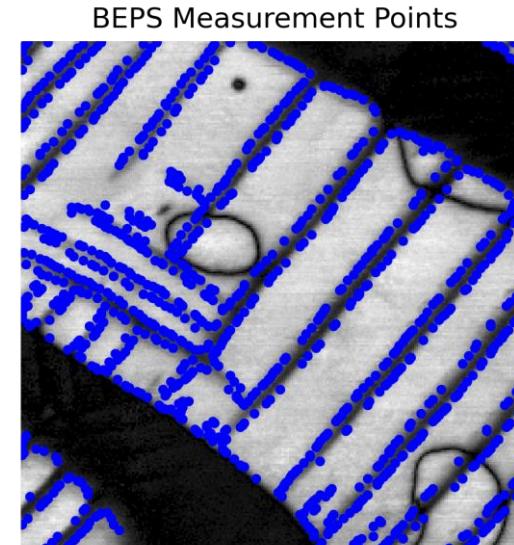
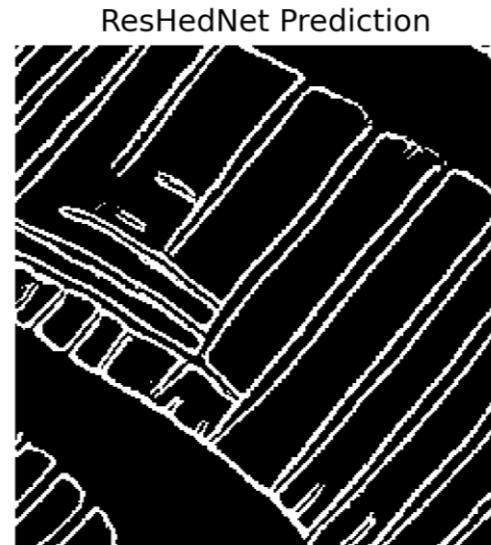
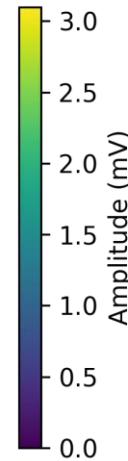
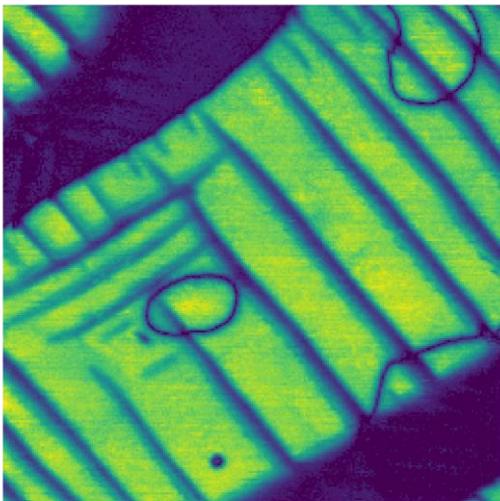
scanning line #56

In []:

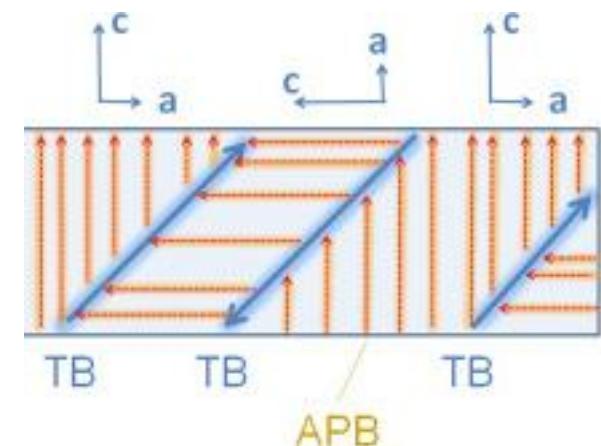
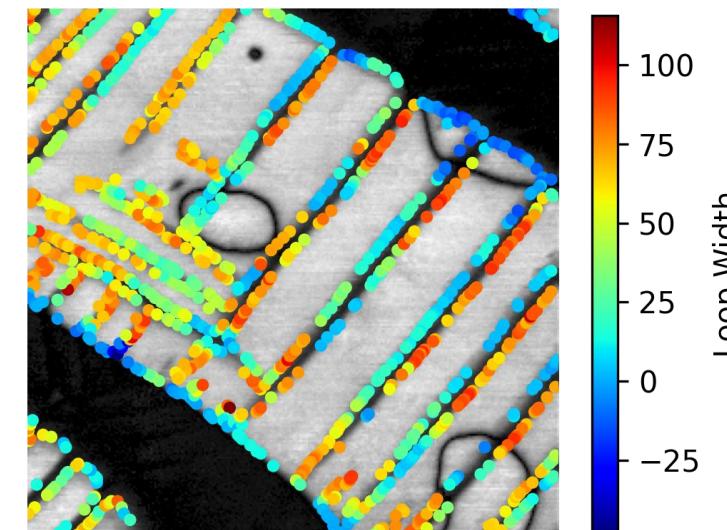
1

In []:

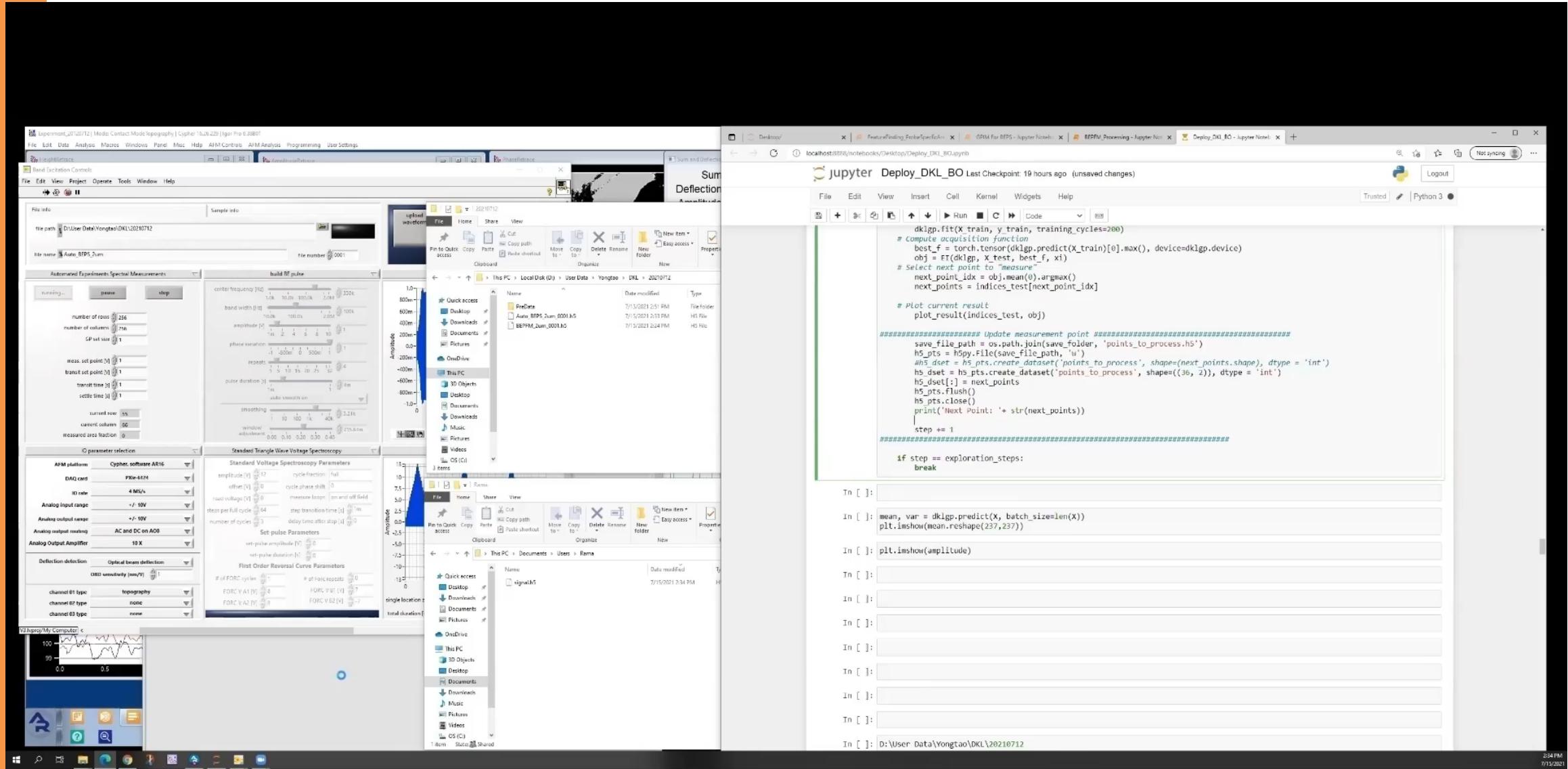
Mapping Activity of Domain Walls



Loop height at ferroelastic walls

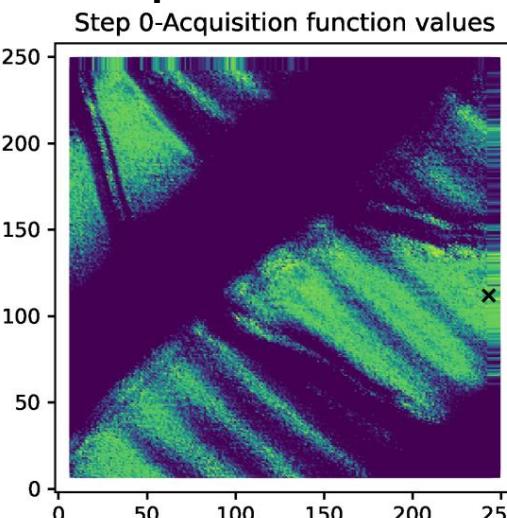
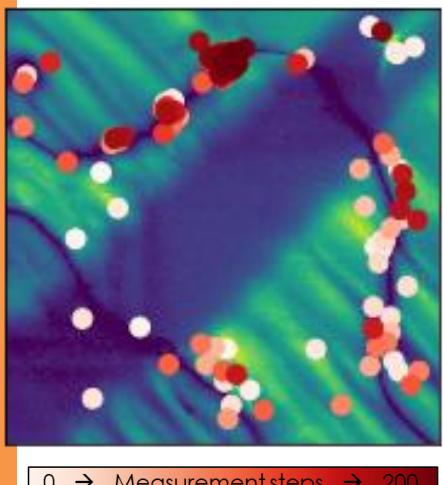


Deep Kernel Learning AE

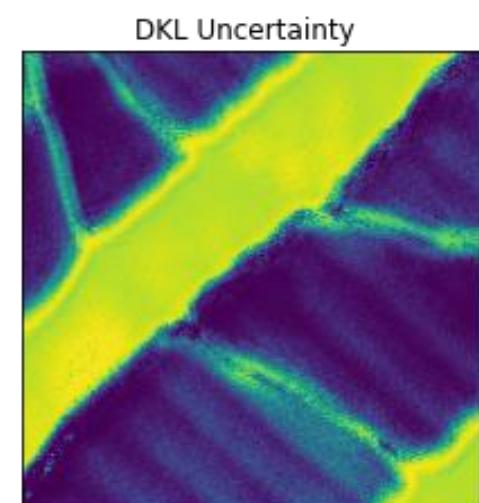
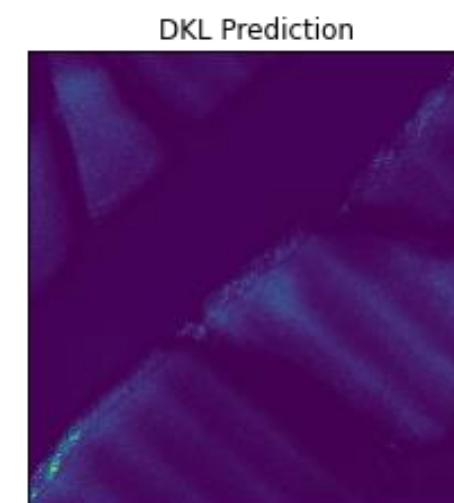
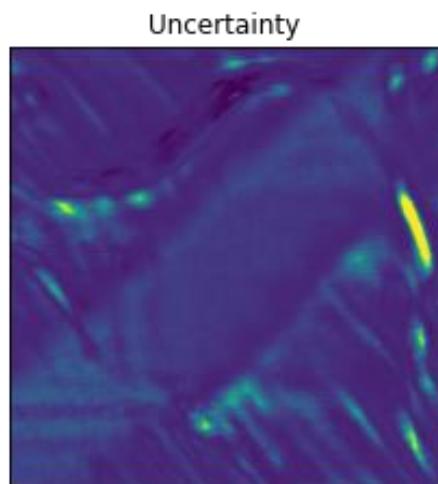
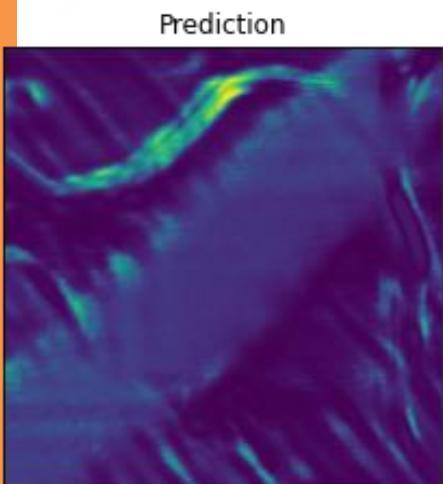
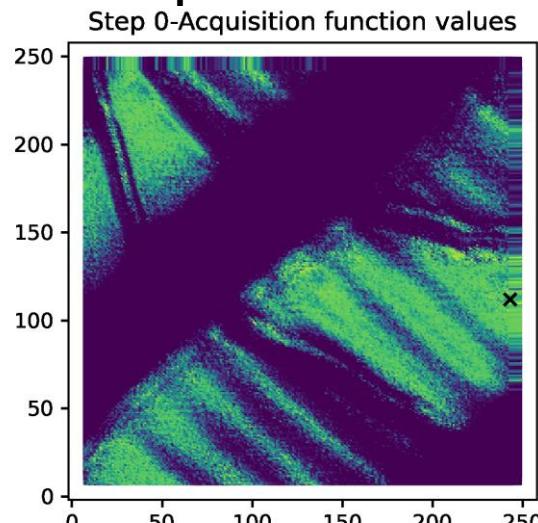
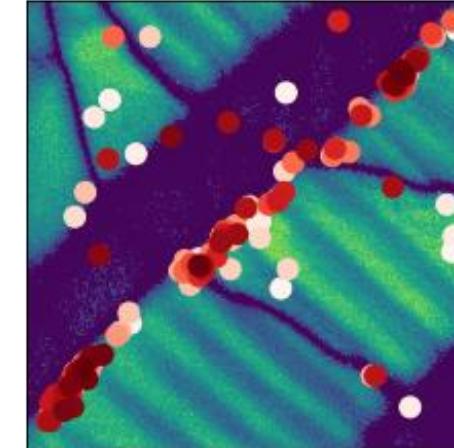


Deep Kernel Learning SPM

Guided by: On field loop area

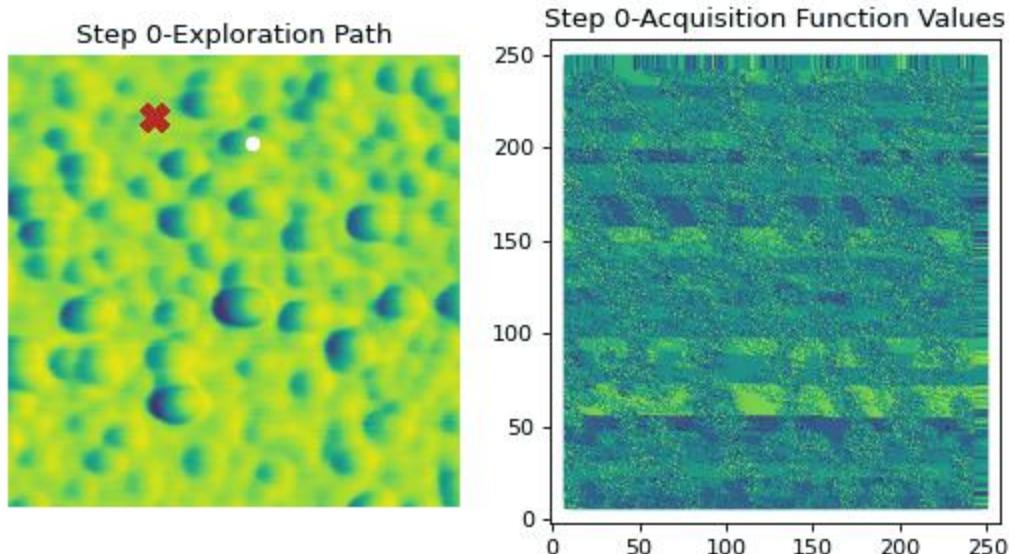
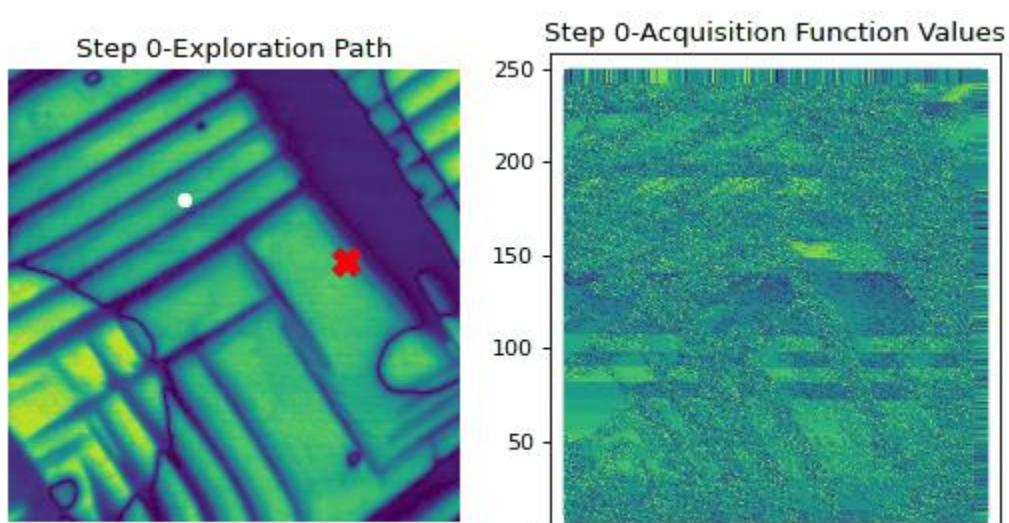
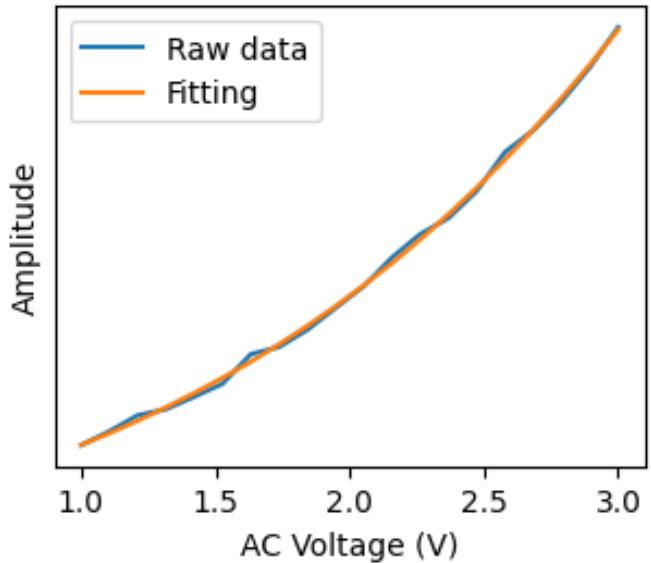


Guided by: Off field loop area



- Large loop opening corresponding 180° domain walls
- This behavior can be attributed to the large polarization mobility of 180° walls

Why human in the loop?



- 200-step automated experiment
- PFM amplitude was used as structure ima
- V_{AC} sweep curve at each location was fitte $y = Ax^3 + Bx^2 + Cx$
- A, B, C, and A/B were used as the target function to guide DKL- V_{AC} measurement.

The methodologies of classical ML (hyperparameter optimization, cross-validation) are rarely applicable for active learning!

The dance of policies and rewards

Rewards and objectives:

- What is our (hierarchical) objective?
- Can we define reward(s)?

Inferential biases:

- What do we know before the experiment?
- What do we (hope to) learn after the experiment?

Experiment planning – policies and values

- How do we plan experiment in advance (policies or values based on rewards)?
- Can we ascribe value to certain steps?
- Do we change our policies during experiment?