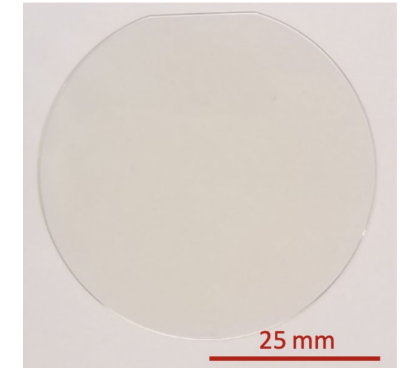
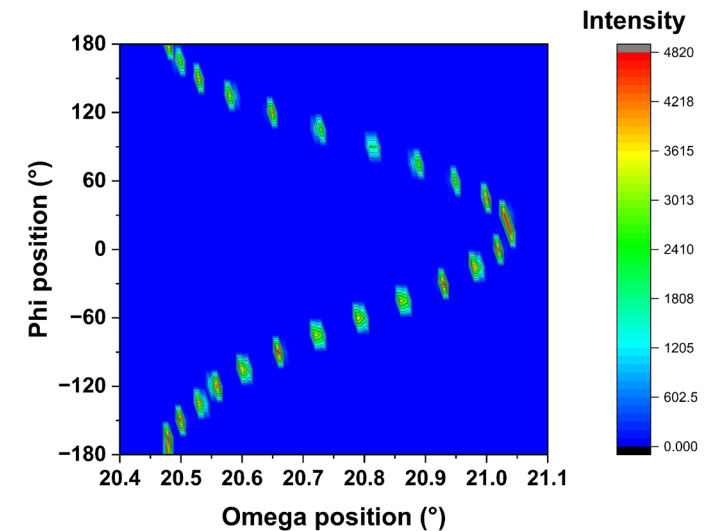
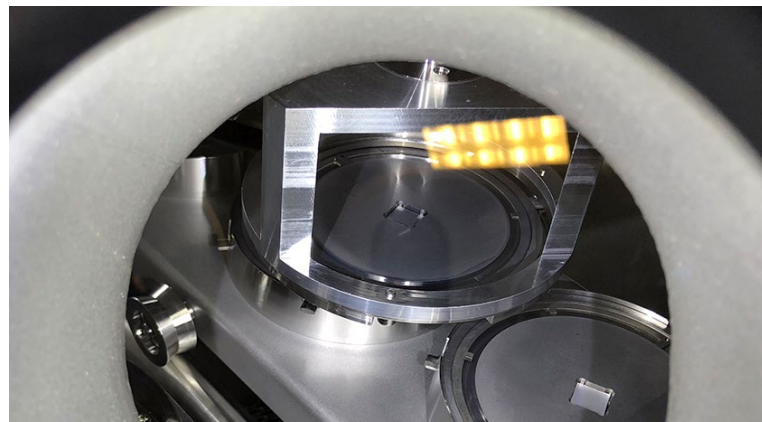
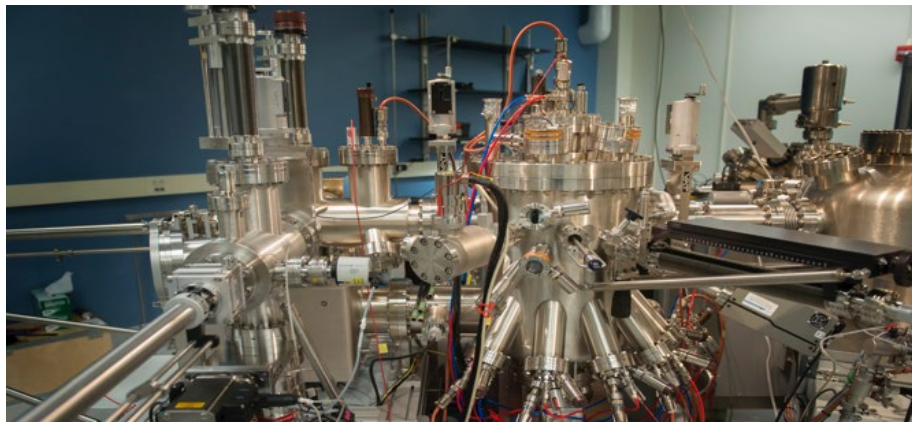


Data Science Activities in the 2DCC-MIP



Joan Redwing, Anthony Richardella, Konrad Hilse, Kevin Dressler, Wesley Reinhart, Nitin Samarth, Vincent Crespi
2D Crystal Consortium (2DCC), Pennsylvania State University





2D Crystal Consortium
NSF Materials Innovation Platform

A national user facility focused on advancing the synthesis of 2D layered chalcogenides

User Facilities

Thin Films and *in situ* Characterization

- MBE, MOCVD and CVD/epi graphene

Bulk Crystal Growth

- CVT, Flux growth, Bridgman, Float Zone

Theory, Simulation and Data Science

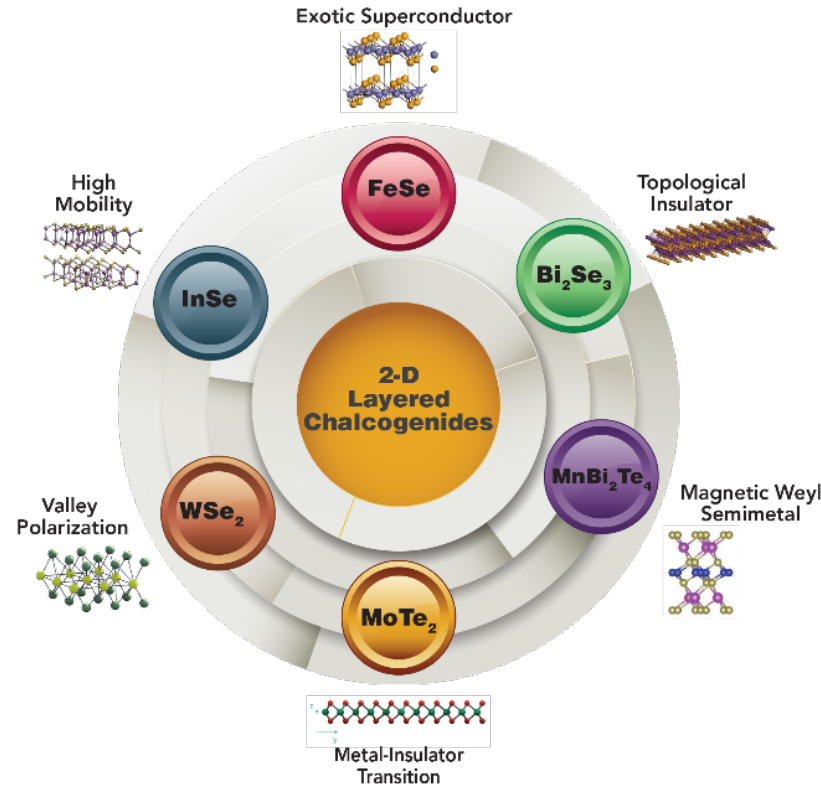
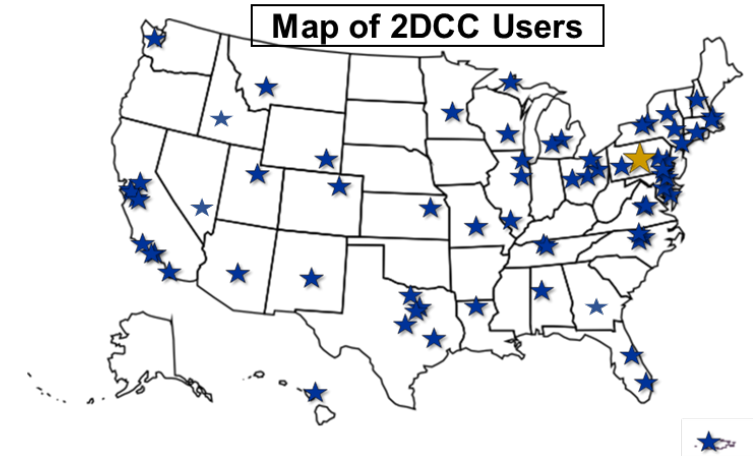
- DFT, ReaxFF-based MD, Phase-field
- LiST database

In-House Expertise



User Program

Map of 2DCC Users



Workshops, Webinars and Hands-on Training

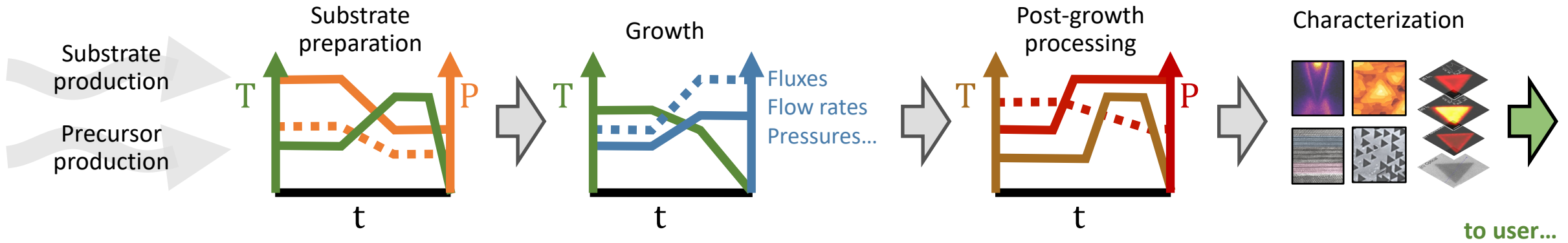


For more information: www.2dccmip.org

In synthesis, every sample has a history...



- Series of processes in time with parameters and uncertainties at every step.



Theorist: "Every sample is identical"

Experimentalist: "Every sample is unique"

- Important to capture the histories, especially the differences, across samples as well as "failed" experiments.

Lifetime Sample Tracking - LiST



- Infrastructure for the capture, curation, and analysis of materials growth and characterization data
 - Sample growth and characterization data > 16,000 bulk and thin film samples
 - Growth recipes/parameters imported directly from the synthesis tool
 - Automated pipelines to collect data wherever possible

The screenshot displays the LiST software interface. On the left, a periodic table is shown with various elements highlighted in different colors. The main interface is divided into several sections:

- CHOOSE PROJECT:** A dropdown menu showing the selected project 'S0026'.
- Sample Activity:** A list of activities for the project, including 'Communication', 'Samples', 'Preparation(MCV1) (12/03/2018)', 'MOCVD (12/03/2018)', 'Raman (12/03/2018)', 'AFM (12/04/2018)', 'Split (12/17/2018)', and 'Raman (12/17/2018)'. The 'MOCVD (12/03/2018)' activity is selected.
- Processing type:** 'Synthesis'.
- Visibility:** 'User'.
- Created by:** 'Mikhail Chubarov'.
- Date:** '12/03/2018'.
- Instrument:** 'MCV1 - MOCVD1'.
- Runtime (h):** '1.5'.
- Favorites:** 'Select a Fav...'
- Material:** 'MoS2'.
- Precursors:** A table listing precursors and their usage.
- Recipe:** A table listing the steps of the synthesis process.

Precursor	Used	Temperature [°C]	Temperature [K]
Mo(CO) ₆	Yes	10	283.15
I: W(CO) ₆ (diluted)	No		
II: W(CO) ₆	No		
DET	No		
H ₂ S	Yes		
H ₂ Se	No		

Step	Time [sec]	Temperature [°C]	Pressure [Torr]	Metal Injector H ₂ [sccm]	Chalcogen Injec [sccm]		
Ramp up	17	1000	50	1000	1000		
Pre growth ann	10	1000	50	1000	1000	2500	80
Growth	15	1000	50	970	600	2500	80

- Available to 2DCC-MIP users to **view their own projects and access data on their samples.**
- Researchers can obtain **direct access** to LiST data via **API**, through a Data Request.
- **Public access to metadata** on all (non-proprietary) LiST samples through the 2DCC website.
- **Data Packages** connected to **data DOIs** support publication and broader data dissemination.

Data Packages



- Paper links to data archived in Penn State ScholarSphere, which links to LiST

nature nanotechnology

Explore content ▾ About the journal ▾ Publish with us ▾

Sign up for alerts 🔔 RSS feed

nature > nature.nanotechnology > articles > article

Article | Published: 27 July 2023

Step engineering for nucleation and domain orientation control in WSe₂ epitaxy on c-plane sapphire

Haoxue Zhu, Nadire Nayir, Tanushree H. Choudhury, Anushka Bansal, Benjamin Huet, Kurvan Zhang, Alexander A. Puzetky, Saibhaneendra Bachu, Krystal York, Thomas V. Mc Knight, Nicholas Trainor, Aarvan Oberoi, Ke Wang, Santarshi Das, Robert A. Makin, Steven M. Durbin, Shenzhi Huang, Nasim Alem, Vincent H. Crespi, Adri C. T. van Duin & Joan M. Redwing

Nature Nanotechnology (2023) | [Cite this article](#)

930 Accesses | 26 Altmetric | Metrics

Abstract

Epitaxial growth of two-dimensional transition metal dichalcogenides on sapphire has emerged as a promising route to wafer-scale single-crystal films. Steps on the sapphire act as sites for transition metal dichalcogenide nucleation and can impart a preferred domain orientation, resulting in a substantial reduction in mirror twins. Here we demonstrate control of both the nucleation site and unidirectional growth direction of WSe₂ on c-plane sapphire by metal-organic chemical vapour deposition. The unidirectional orientation is found to be intimately tied to growth conditions via changes in the sapphire surface chemistry that control the step edge location of WSe₂ nucleation, imparting either a 0° or 60° orientation relative to the underlying sapphire lattice. The results provide insight into the role of surface chemistry on transition metal dichalcogenide nucleation and domain alignment and demonstrate the ability to engineer domain orientation over wafer-scale substrates.

Main

00:04:46 | [Two-dimensional \(2D\) transition metal dichalcogenides \(TMDs\) with desired properties are](#)

ScholarSphere Penn State University Libraries

About Help Contact Login

WSe₂ on c-plane sapphire with preferred orientation grown by MOCVD

This data package includes information on sample preparation, growth conditions and characterization results for WSe₂ samples reported in the main body of the manuscript entitled "Step engineering for nucleation and domain orientation control in WSe₂ epitaxy on c-plane sapphire" by H. Zhu, et al. The WSe₂ samples were grown by metalorganic chemical vapor deposition (MOCVD) on 2" c-plane sapphire substrates in the NSF 2D Crystal Consortium facility at Penn State (systems MOCVD1 and MOCVD2). The data includes recipe files from the MOCVD equipment and data files from atomic force microscopy (AFM), field emission scanning electron microscopy (FE-SEM) and room temperature photoluminescence and Raman spectroscopy carried out on the samples. Characterization was typically done at the center of the 2" wafer and near the wafer edge (approximately 5 mm from wafer edge on opposite side to the flat) unless otherwise specified.

All samples grown as part of this study are included in the data package. The sample numbers associated with figures in the main manuscript are as follows:

Figure 1 MCV1-200316A-HZ, MCV1-200814A-HZ, MCV1-200813A-HZ, MCV1-200313A-HZ, MCV1-200312A-HZ Figure 2 MCV2-200914A-TM, 2009004A-TM Figure 3 MCV1-200311A-HZ, MCV1-200310A-HZ Figure 4 MCV2-210723A-KY, MCV1-211005A-TM Figure 5 MCV1-200311A-HZ, MCV1-200316A-HZ, MCV1-200309A-HZ Figure 6 MCV1-200722A-HZ

Additional details are provided in the 2DCC LIST database: <https://m4-2dcc.vmhost.psu.edu/list/data/10.26207/7b1-fs46>

Joan Redwing jmr31@psu.edu 5/25/23

FILES

MCV1-191016A-BH.zip	size: 12 MB mime_type: application/zip date: 2023-05-26 sha256: 3a15882
MCV1-191101A-BH.zip	size: 4.42 KB mime_type: application/zip date: 2023-05-26 sha256: c3be089
MCV1-191107A-HZ.zip	size: 15.3 MB mime_type: application/zip date: 2023-05-26 sha256: 9d6a705
MCV1-191111A-HZ.zip	size: 4.62 KB mime_type: application/zip date: 2023-05-26 sha256: 287e170

METADATA

Work Title WSe₂ on c-plane sapphire with preferred orientation grown by MOCVD

Subtitle Materials Science

Access [OPEN ACCESS](#)

Creators Joan Redwing, Haoxue Zhu, Tanushree Choudhury, Thomas Mc Knight, Nicholas Trainor, Benjamin Huet

Keyword MOCVD, WSe₂, Sapphire

License CC BY 4.0 (Attribution)

Work Type Dataset

Acknowledgments NSF Cooperative Agreement

Publisher 2D Crystal Consortium

Publication Date May 25, 2023

Subject Materials Science

DOI [10.26207/7b1-fs46](https://doi.org/10.26207/7b1-fs46)

Related URLs <https://m4-2dcc.vmhost.psu.edu/list/data/10.26207/7b1-fs46>

Deposited May 26, 2023

VERSIONS

V3	PUBLISHED
V2	PUBLISHED
V1	PUBLISHED

ANALYTICS

134 views since May 26, 2023

2DCC-MIP

Publication

Year: 2023 | Journal: Nature Nanotechnology

Science Driver: Epitaxy of 2D Chalcogenides | Other Science Drivers: Advanced Characterization and Modeling

Title: Step engineering for nucleation and domain orientation control in WSe₂ epitaxy on c-plane sapphire

Description: [Redacted]

Authors: H. Zhu, N. Nayir, T.H. Choudhury, A. Bansal, B. Huet, K. Zhang, A.A. Puzetky, S. Bachu, K. York, T.V. Mc Knight, N. Trainor, A. Oberoi, K. Wang, S. Das, R.A. Makin, S.M. Durbin, S. Huang, N. Alem, V.H. Crespi, A.C.T. van Duin, J.M. Redwing

Citation: H. Zhu, N. Nayir, T.H. Choudhury, A. Bansal, B. Huet, K. Zhang, A.A. Puzetky, S. Bachu, K. York, T.V. Mc Knight, N. Trainor, A. Oberoi, K. Wang, S. Das, R.A. Makin, S.M. Durbin, S. Huang, N. Alem, V.H. Crespi, A.C.T. van Duin, J.M. Redwing, "Step engineering for nucleation and domain orientation control in WSe₂ epitaxy on c-plane sapphire", Nature Nanotechnology (2023). [10.1038/s41565-023-01456-6](https://doi.org/10.1038/s41565-023-01456-6)

DOI: [10.1038/s41565-023-01456-6](https://doi.org/10.1038/s41565-023-01456-6)

Publication Type: [Redacted] | Associated Projects: [Redacted]

In-House: [Redacted] | Enter ID in field to the right

Materials Available Data Synthesis Techniques Available Characterization Data

x Any Any Any

RESET SEARCH CRITERIA SEARCH

Search in Title, DOI and Description

Show 10 entries

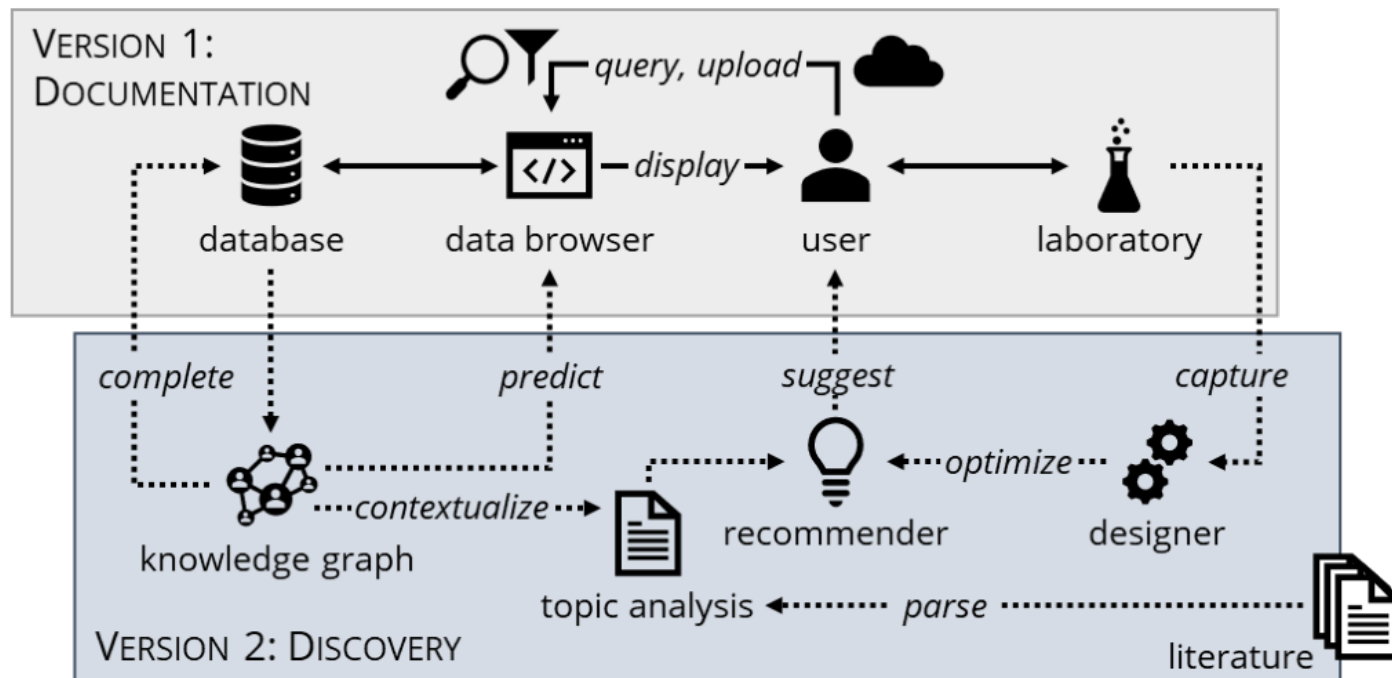
Label	2DCC Data DOI	Project	Publication DOI	Available Data	Type
Young Modulus		L_ReaxFF_2D_Toolbox		Theory Data	Snapshot
WSe2 on c-plane sapphire with preferred orientation grown by MOCVD	10.26207/f7b1-fs46		10.1038/s41565-023-01456-6	87 Samples	Snapshot - Published
Wafer-scale MOCVD TMD films used for 3D Monolithic Integration	10.26207/khwb-rr73			38 Samples	Snapshot - Published
Structures of defect pairs		L_SIM_BN_MoS2_1	10.1103/PhysRevB.99.155430	Theory Data	Snapshot
Potatoe Stamp Concept		L_ReaxFF_2D_Toolbox	10.1021/acs.jpcc.8b02991	Theory Data	Snapshot
MoSe2-WSe2 in-plane heterostructures for exciton confinement				21 Samples	Snapshot - Published
InSe on 3 Inch Si(111) Wafer-Scale Synthesis with Combinatorial Growth Approach by MBE	10.26207/qakv-p610			13 Samples	Snapshot - Published

Showing 1 to 7 of 7 entries

Previous 1 Next



LiST 2.0: From Data Store to Knowledge Graph



1.0: Documentation
↓
2.0: Discovery

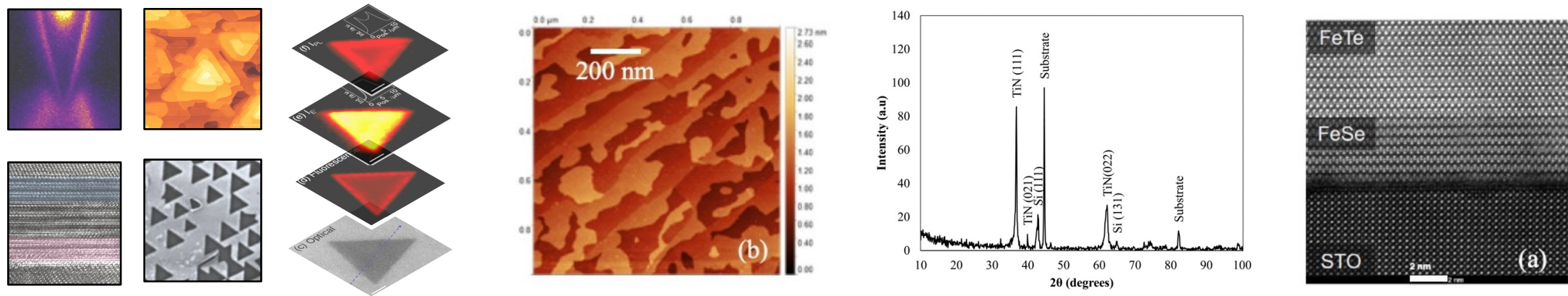
LiST 1.0 is a **traditional database** whose consistent schema are straightforward to query, synthesize, and maintain, but it can only map explicit relationships between data.

List 2.0's **knowledge graph** will contain both a data store and an ontology describing the relationships between entities and will be able to infer relationships not explicitly provided.



Wes Reinhart
Data Science Lead

Why should materials researchers care about computer vision?



- In the data science sense, “images” are any data with regular spatial structure. Images thus are an abundant data modality for 2DCC.
- Each image contains more data than other common modalities like text or scalars.
- Humans under-utilize image data (in science) as we can’t study them pixel-by-pixel.

Thus, computer vision will be a key component of materials informatics.

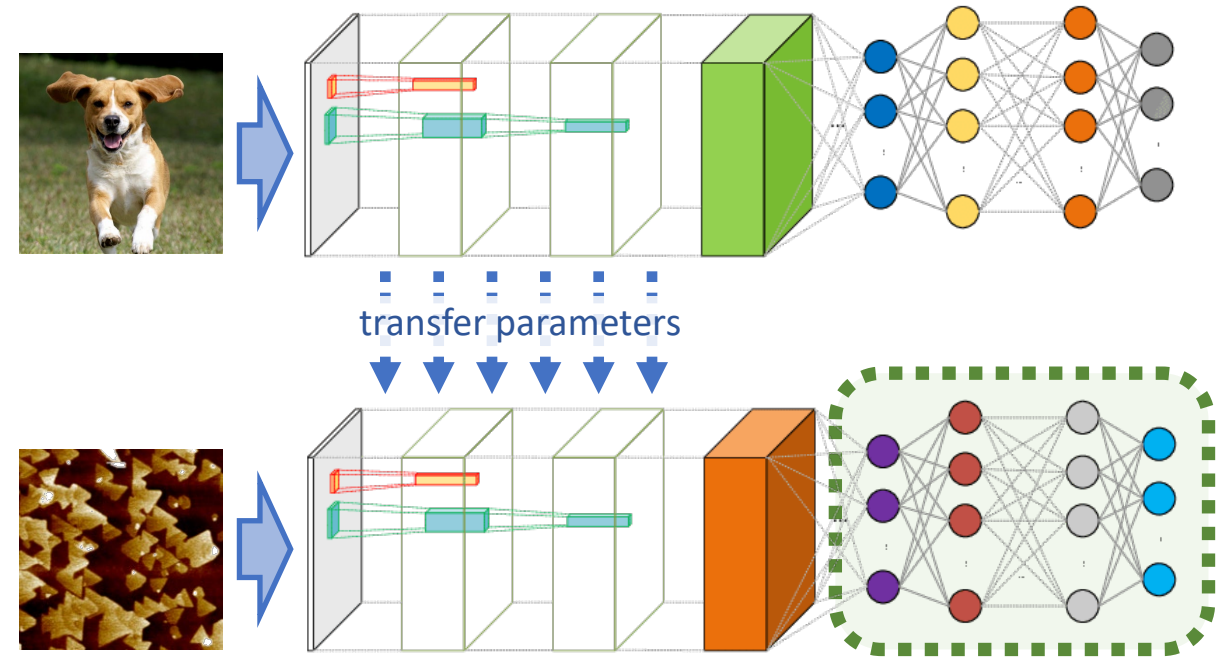
Pretrained Convolutional Neural Networks

Convolutional Neural Networks learn abstractions for patterns in large image collections, such as classifying “dog” versus “cat”. They can have billions of learned parameters and leverage the internet for data.

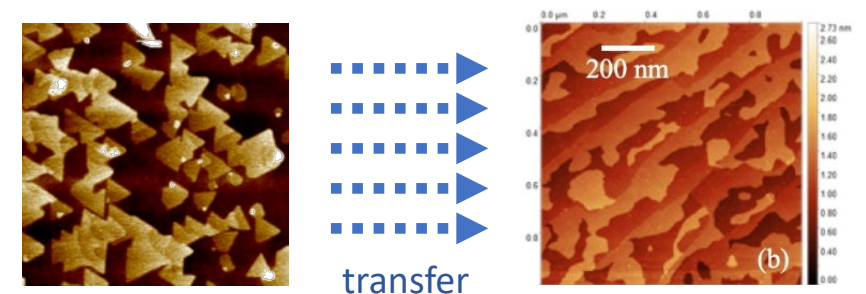
Transfer learning can retrain just a few of these weights to **specialize this highly trained classifier to a new domain** that has a more modest image collection.

More recently, we have also been leveraging “feature extraction” with pretrained networks to evaluate patterns in micrographs without any network retraining.

Transfer learning can leverage pattern recognition developed in one (non-materials) context into new (materials research) contexts, and also between different materials datasets (e.g. MOCVD→MBE).



Lemley, Bazrafkan, and Corcoran (2017) MAICS



Recent examples of transfer learning:

Monolayer coverage of WSe₂

We perform systematic testing of different ML methods for estimating a scalar quantity from 221 AFM. Top-performing models have $R^2 = 0.98$.

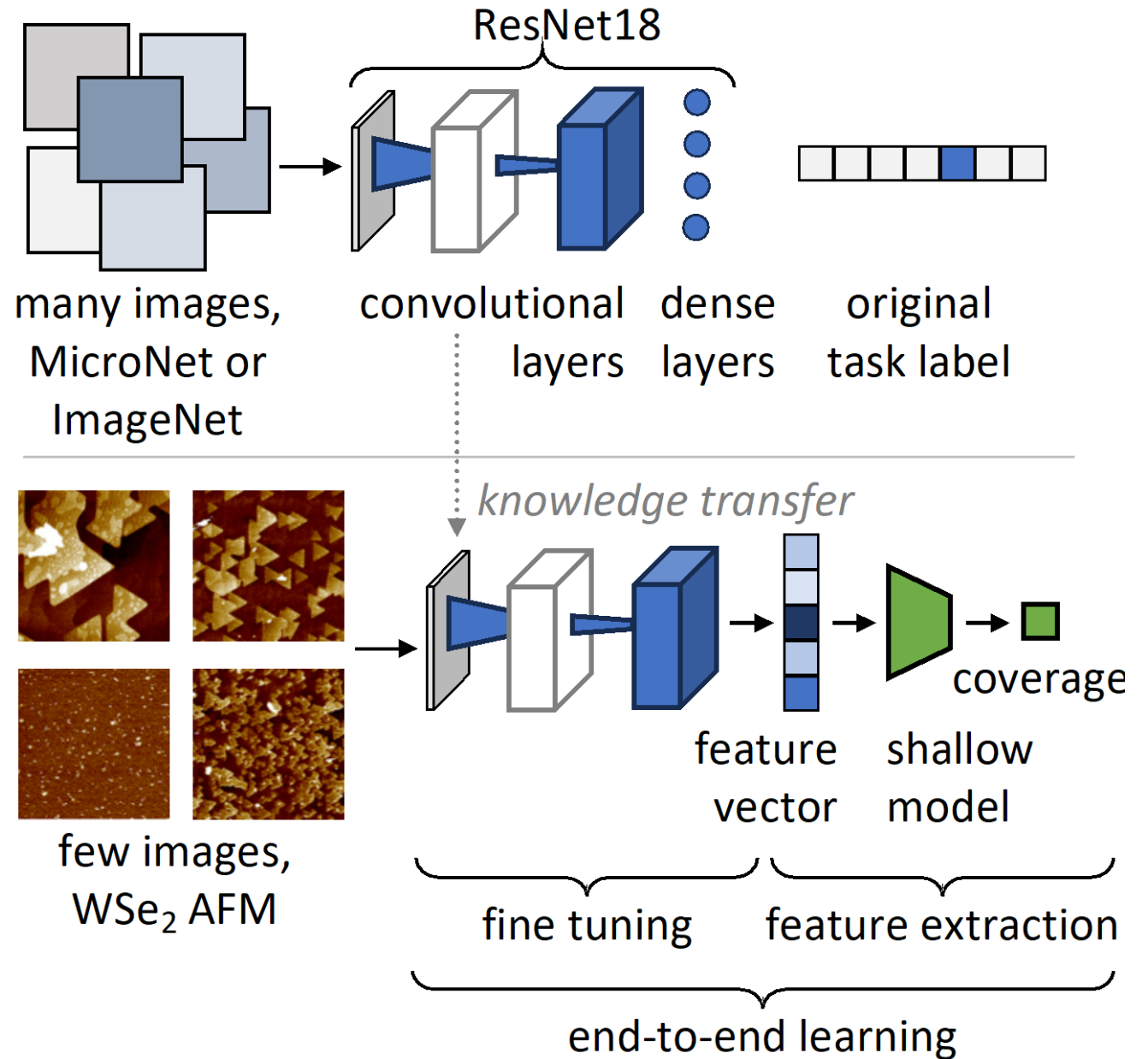
I.A. Moses, C. Wu, W.F. Reinhart, *Materials Today Advances*, 2024, 22, 100438.

Classification of MoS₂ growth conditions

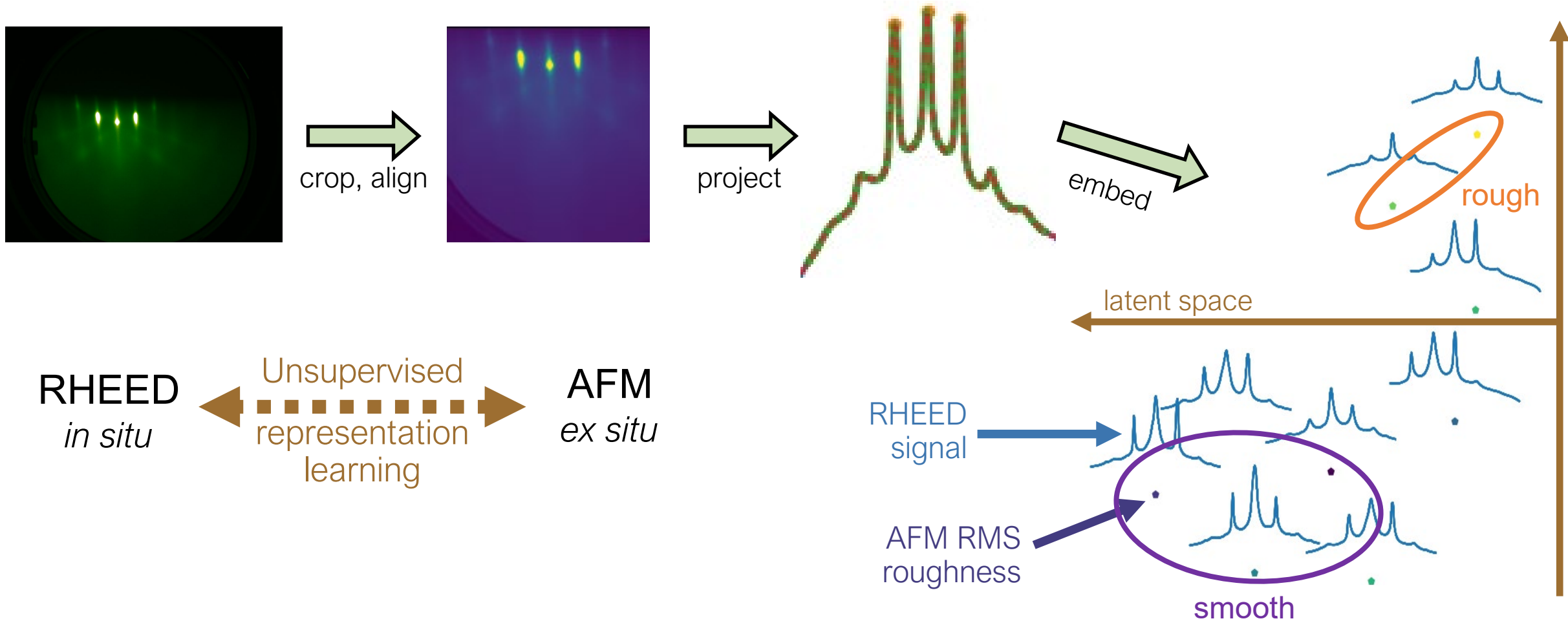
We investigate many different classification methods to retroactively “predict” the growth conditions from 262 AFM. Best $R^2 = 0.70$.

I.A. Moses, W.F. Reinhart, *Materials Characterization*, 2024, 209, 113701.

This appears to be an effective strategy for producing accurate models with few data.



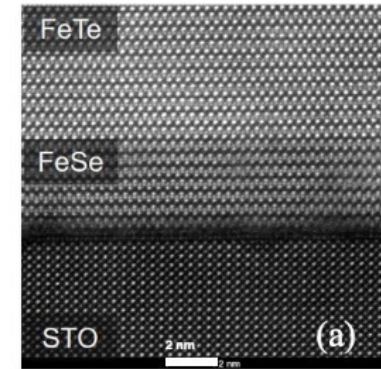
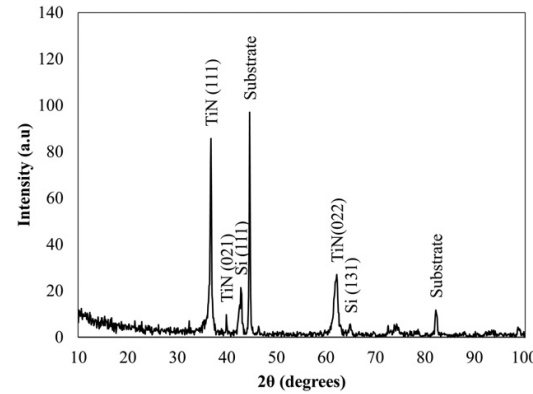
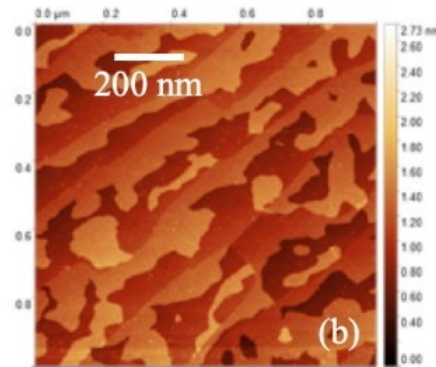
Reflection high-energy electron diffraction (RHEED)



Correlating signals between instruments unlocks new predictive capabilities, especially if *in situ* measurements acquired during growth can predict *ex situ* ones measuring growth outcomes.

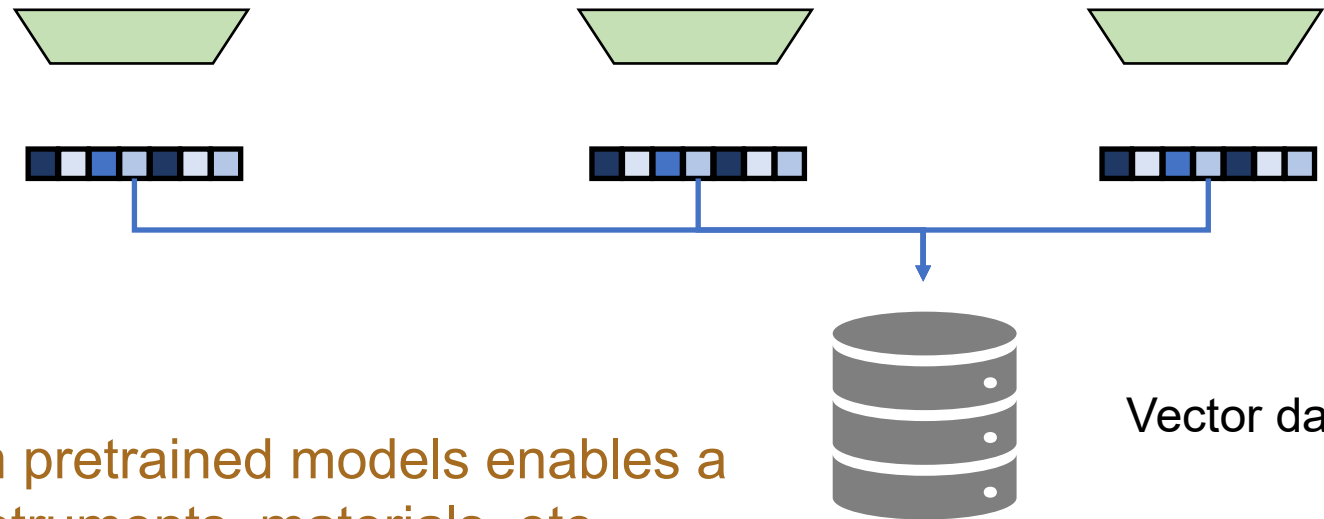
Vector database for multi-modal data

Big, multi-modal data source (i.e., LiST)
 – how to compare different modalities?



Pretrained computer vision models

Everything is a vector!



Leveraging feature extraction from pretrained models enables a common representation across instruments, materials, etc. Once in vector form, standard analyses can be performed.

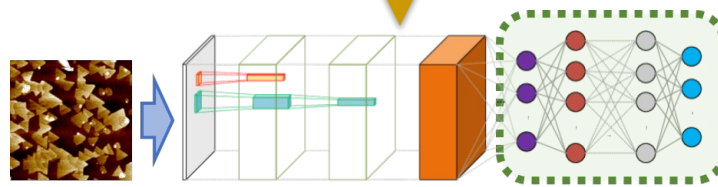
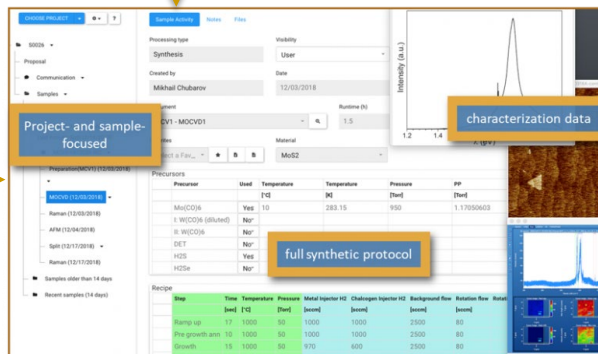


Advancing 2D materials synthesis through data science

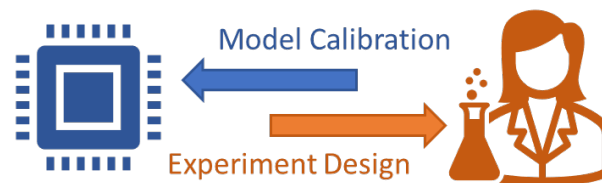
In situ and post-growth characterization data



Source information
Sample prep
Process recipes
Lab sensor data
User notes



AI assistant to improve human decision making and experimental design



LiST 1.0

- Automated data capture and organization
- Facilitates **storage and sharing**

LiST 2.0

- Computer vision tools and vectorization
- Knowledge graph generation
- Facilitates **analysis & correlations**

LiST 3.0

- Bayesian design of experiments
- Active learning
- Facilitates **human-AI collaboration**

2DCC Data Science Team



LiST 1.0

LiST 2.0



Konrad Hilse
Full Stack Dev



Kevin Dressler
Operations Manager



Vin Crespi
Theory, Simulation, Data
Science Facility Lead



Wes Reinhart
Data Science Lead



Isaiah Moses
Postdoctoral Scholar



Tool Integration

Anthony Richardella
MBE Research Prof