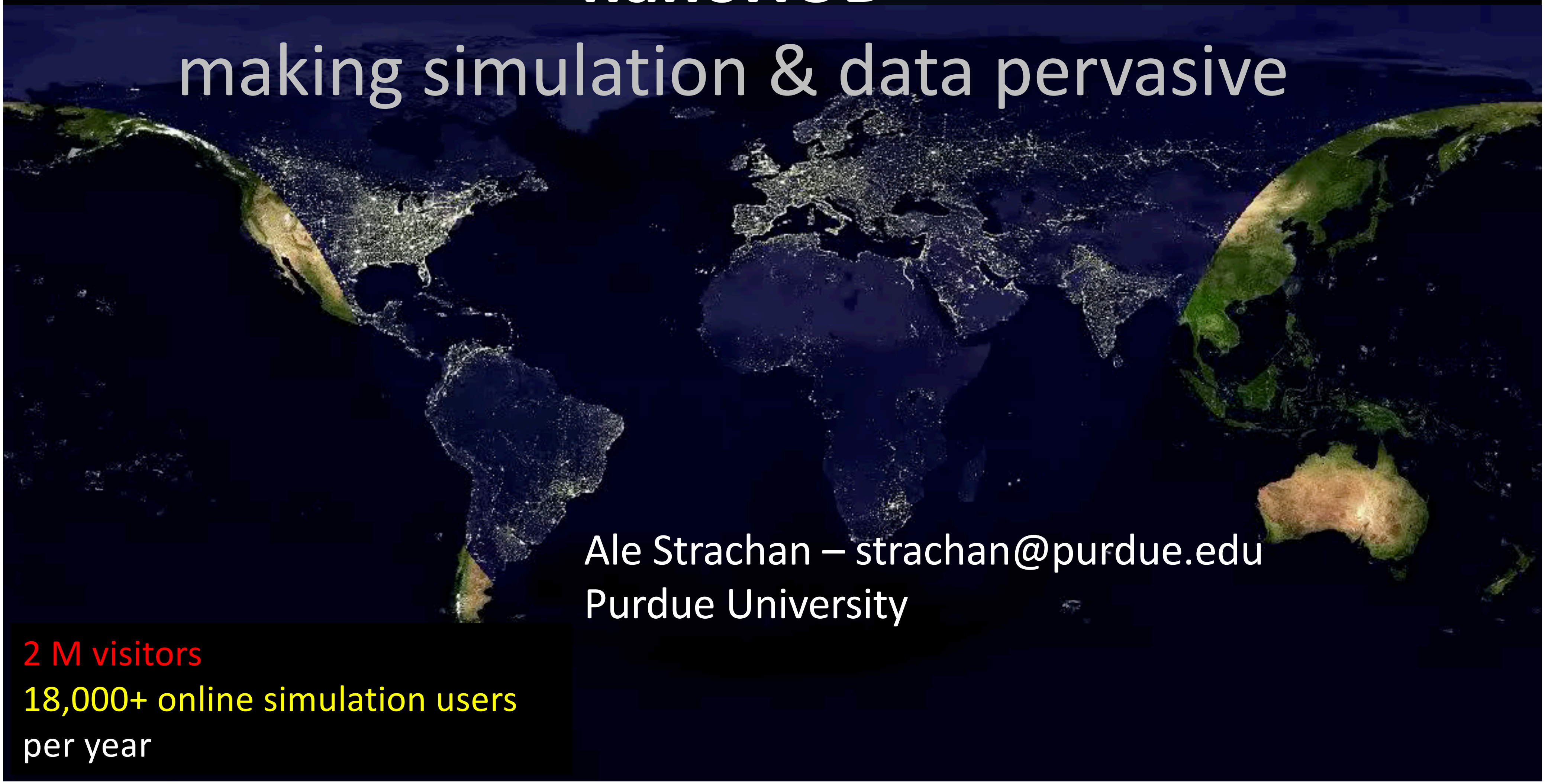
nanoHUB

making simulation & data pervasive



Ale Strachan – strachan@purdue.edu
Purdue University

2 M visitors
18,000+ online simulation users
per year

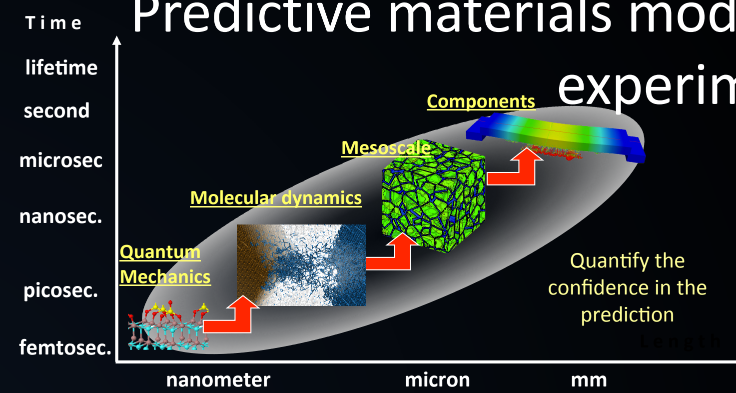

BARRON'S Interests Magazine Data Advisor Penta

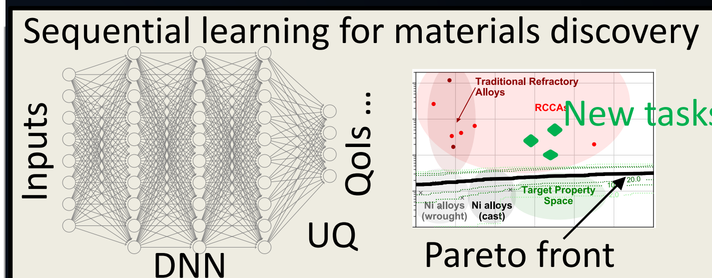
TECHNOLOGY OTHER VOICES

3 Technologies That Could Create Trillion-Dollar Markets Over the Next Decade

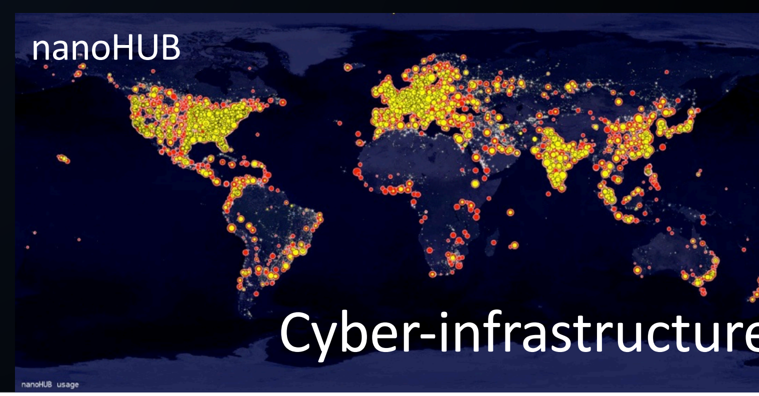
By Greg Satell Updated April 21, 2019 / Original February 17, 2019

Predictive materials models & experiments

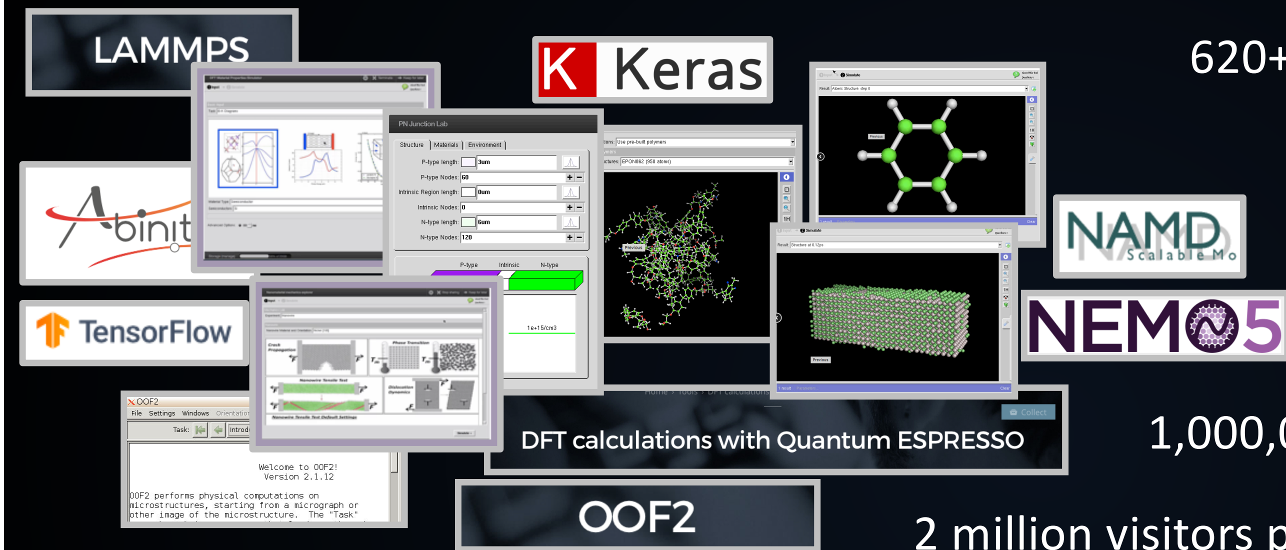




Data science & machine learning



nanoHUB: a community-driven resource



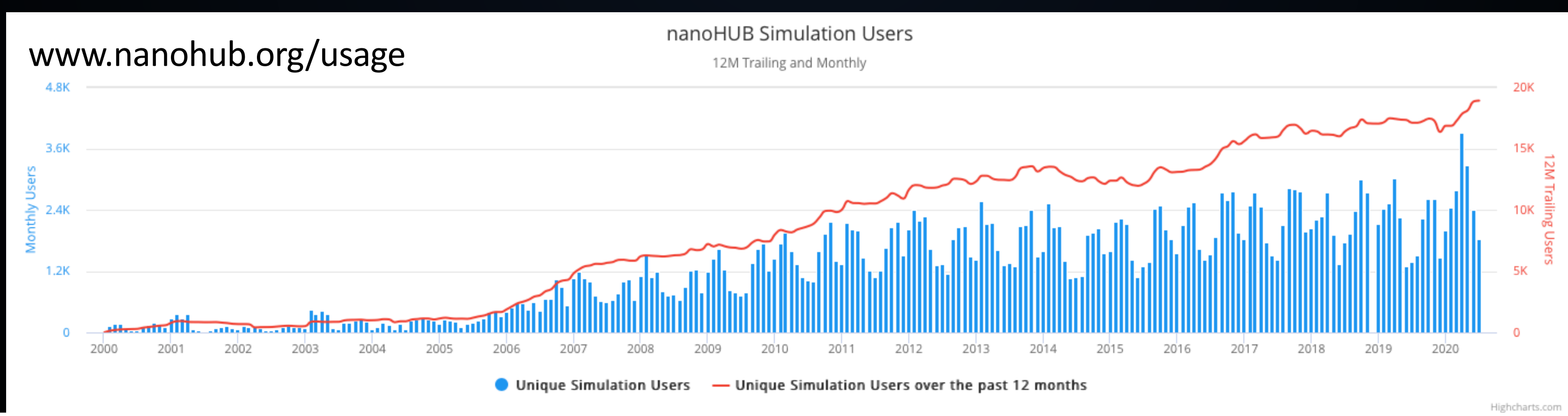
620+ apps & tools

1,800+ contributors

~18,000+ annual simulation users

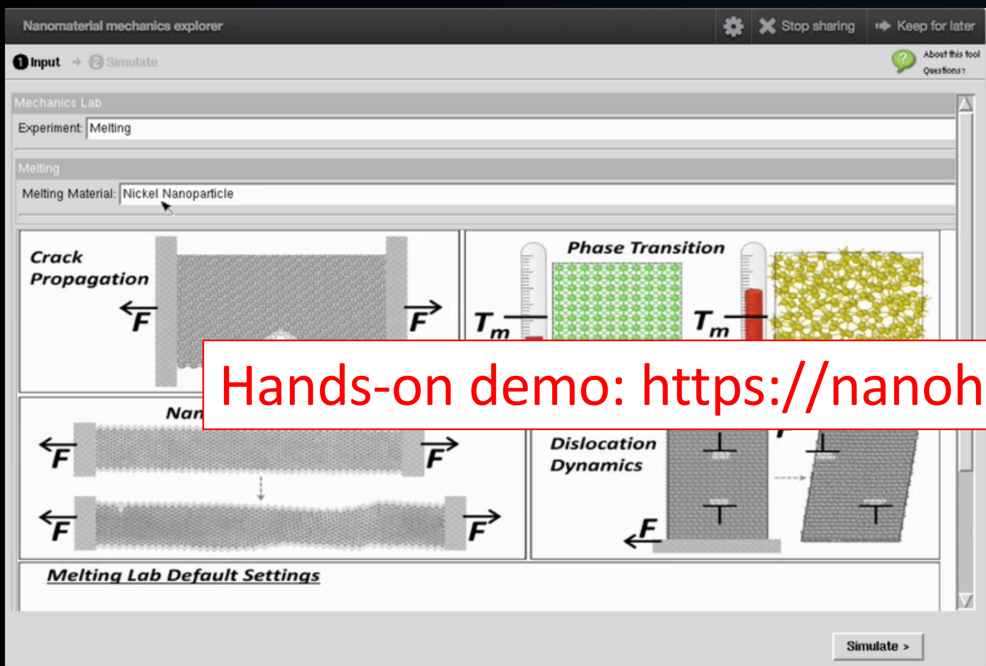
1,000,000+ simulations per year

2 million visitors per year



Apps connected to powerful research codes

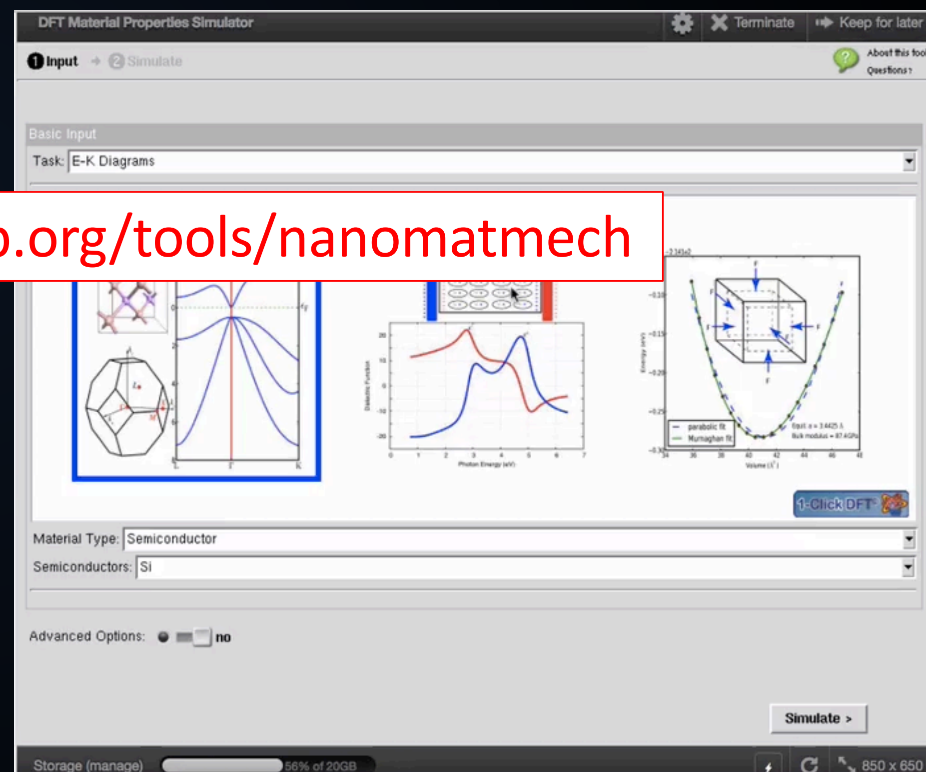
Designed for end users – instructors, students, domain experts



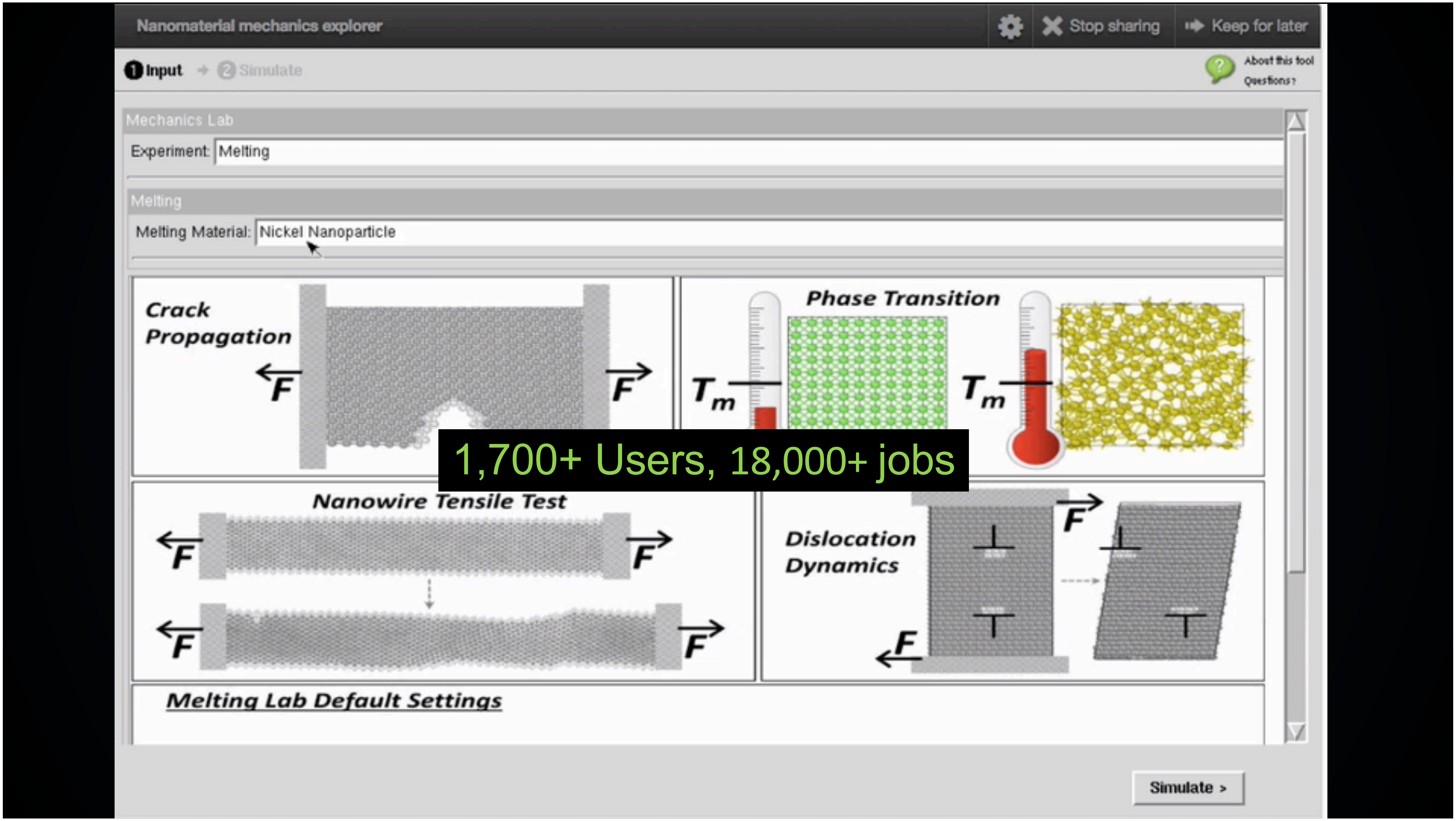
https://nanohub.org/tools/nanomatmech
Powered by LAMMPS

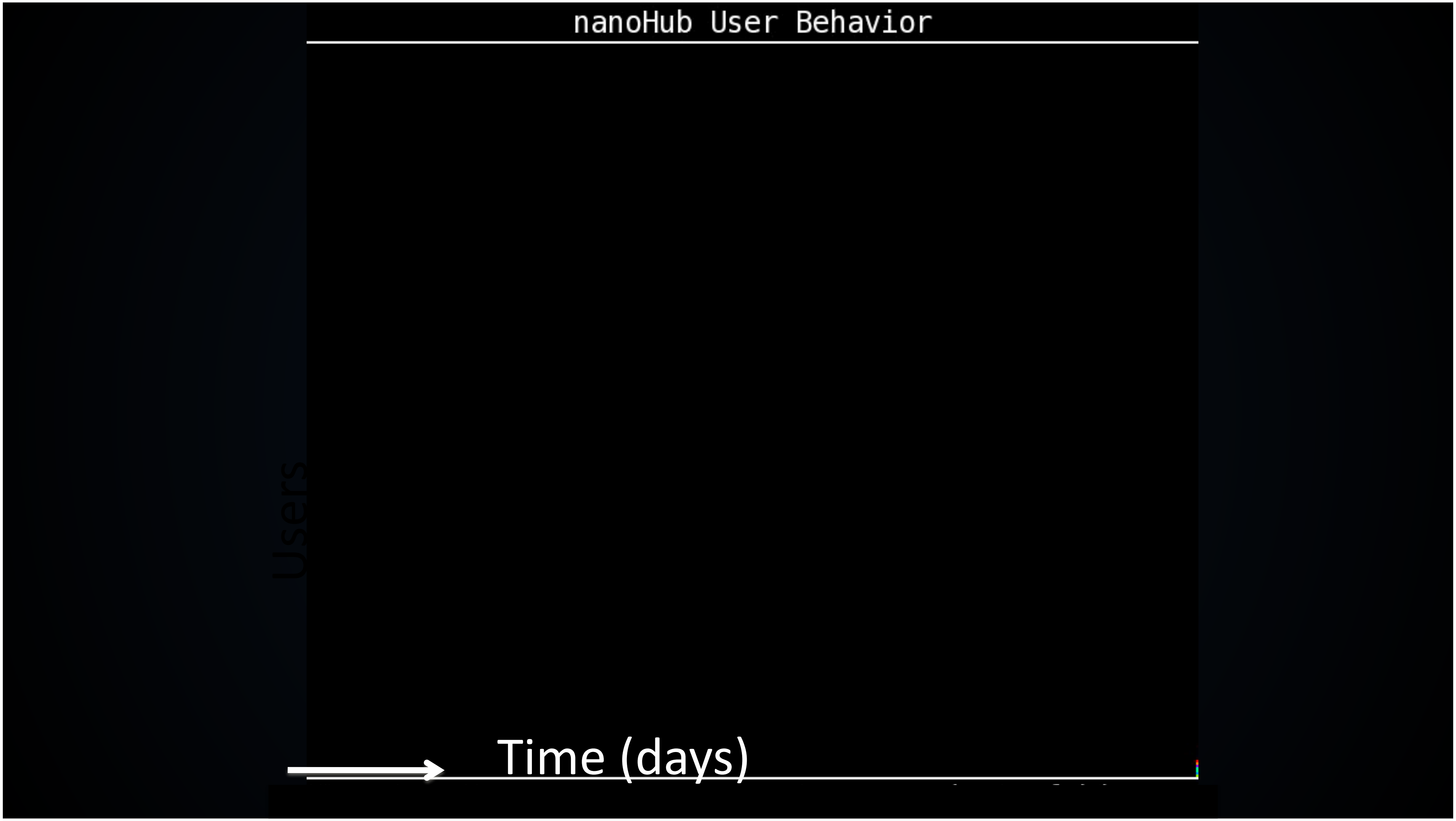
Hands-on demo: https://nanohub.org/tools/nanomatmech

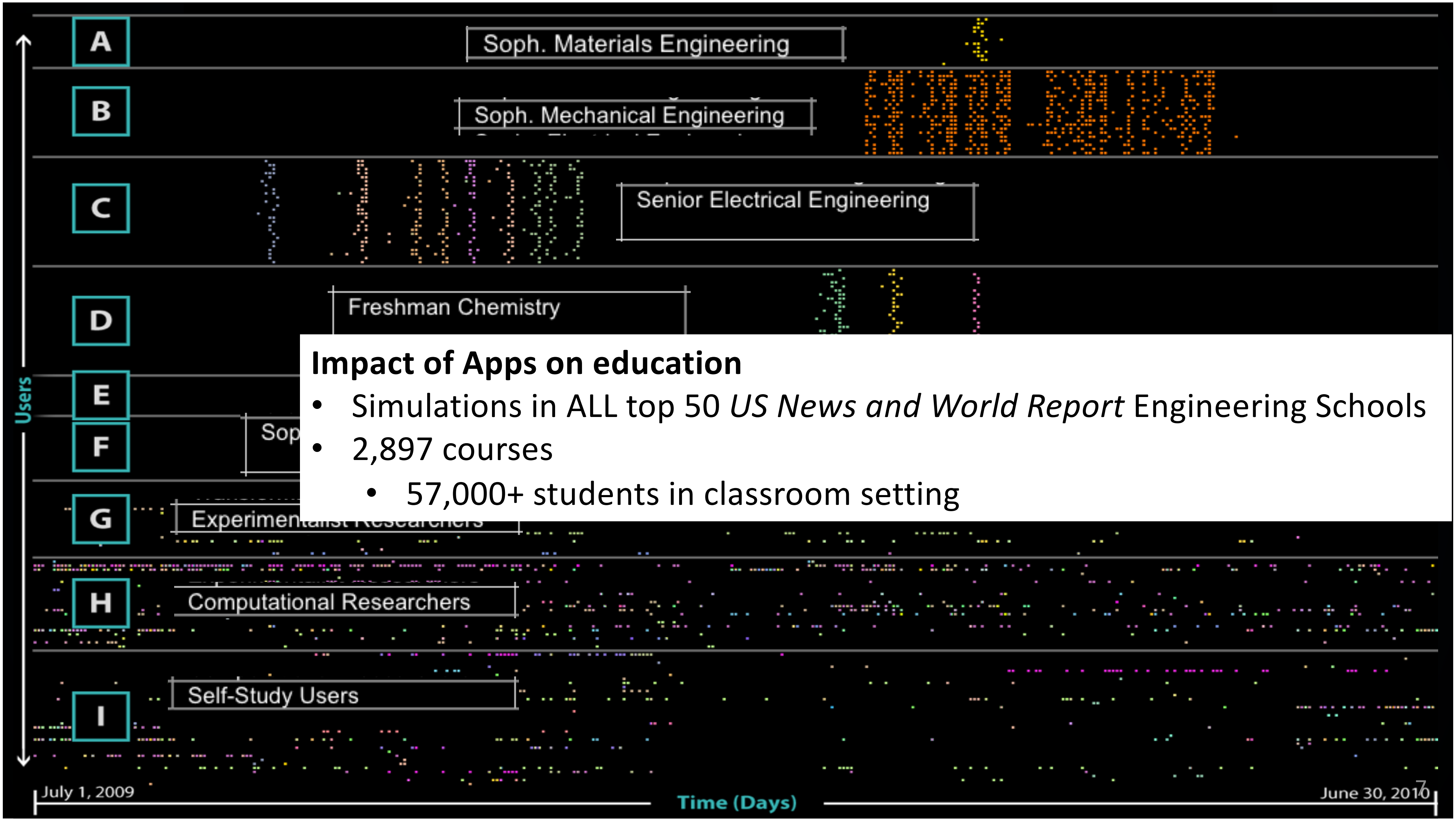
https://nanohub.org/tools/dftmatprop
Powered by Quantum Espresso



Reeve, Guzman, Alzate-Vargas, Haley, Liao, & Strachan, MRS Advances, 1-16. doi:10.1557/adv.2019.287

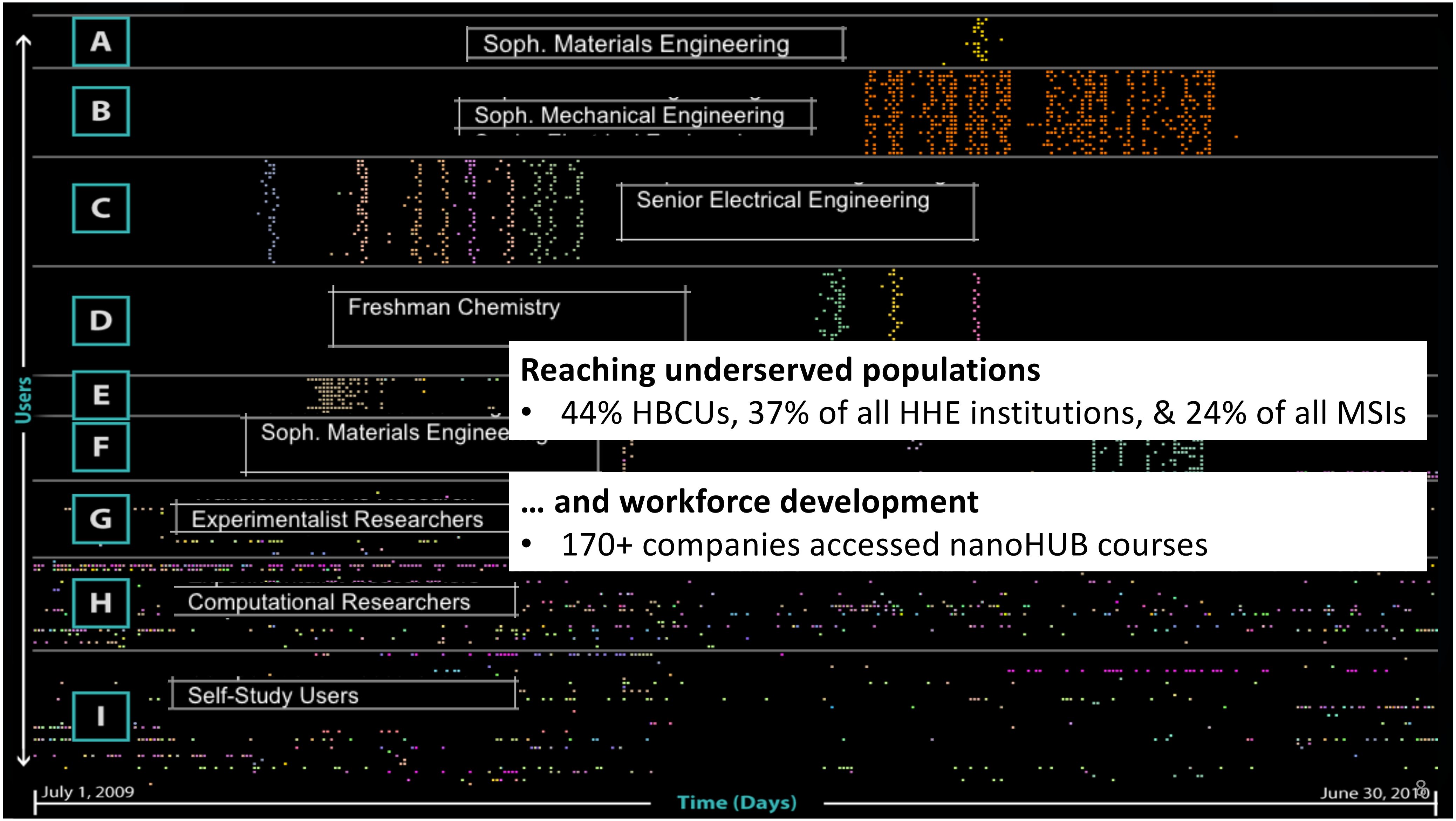

Nanomaterial mechanics explorer
Stop sharing
Keep for later
1 Input
2 Simulate
About this tool
Questions?
Mechanics Lab
Experiment: Melting
Melting
Melting Material: Nickel Nanoparticle
Crack Propagation
Phase Transition
Nanowire Tensile Test
Dislocation Dynamics
Melting Lab Default Settings
Simulate >
1,700+ Users, 18,000+ jobs


nanoHub User Behavior
Users
Time (days)



Impact of Apps on education

- Simulations in ALL top 50 *US News and World Report* Engineering Schools
- 2,897 courses
  - 57,000+ students in classroom setting



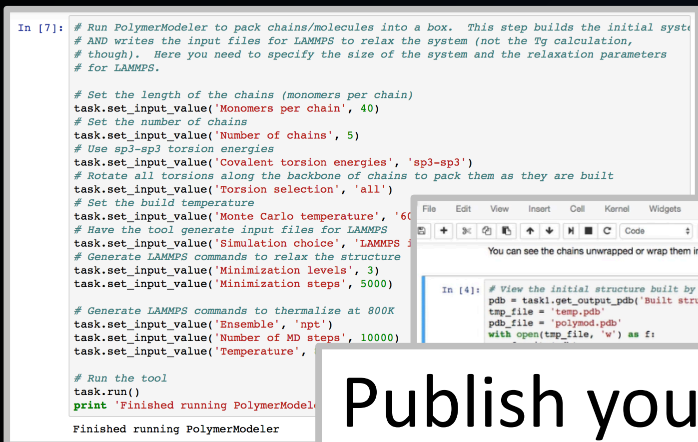
Reaching underserved populations

- 44% HBCUs, 37% of all HHE institutions, & 24% of all MSIs

… and workforce development

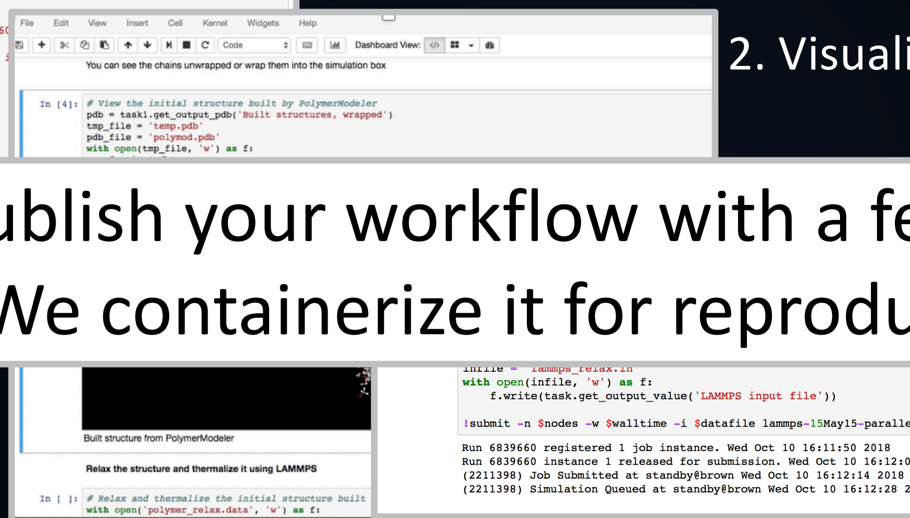
- 170+ companies accessed nanoHUB courses

Jupyter: end to end scientific workflows



1. Use *Polymer Modeler* to create amorphous polymer system

2. Visualize structure

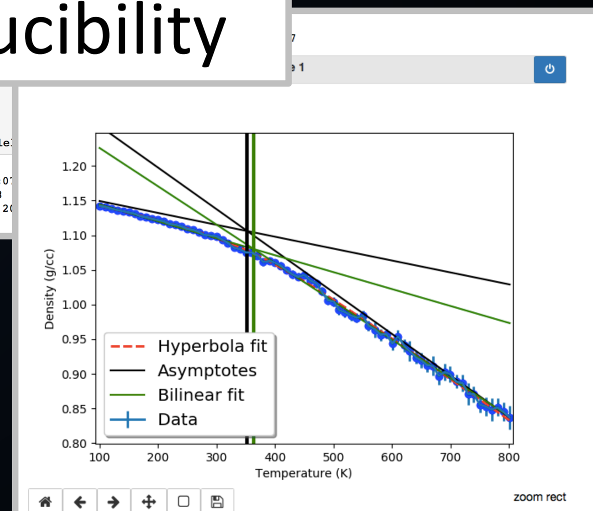


Publish your workflow with a few clicks

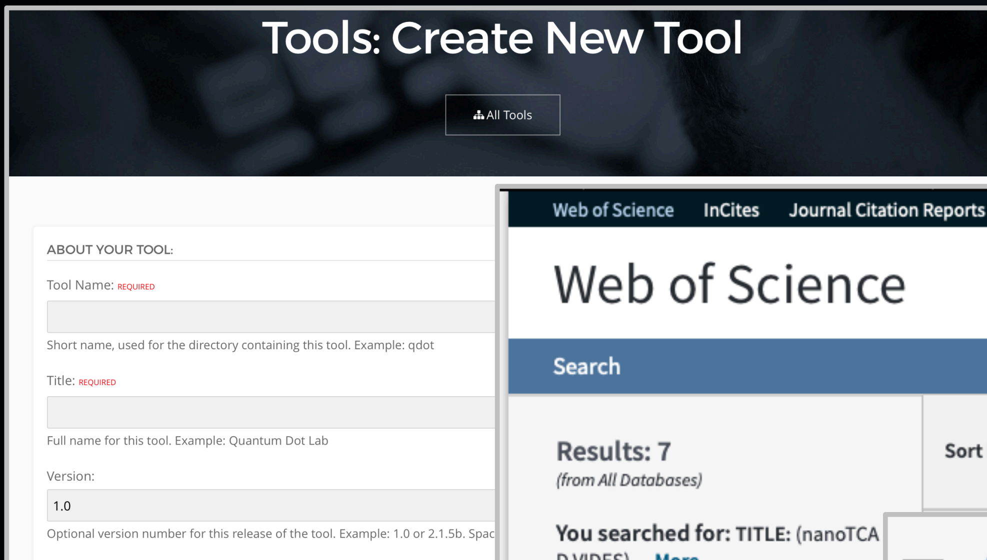
We containerize it for reproducibility

3. MD simulations using *LAMMPS* on HPC resources

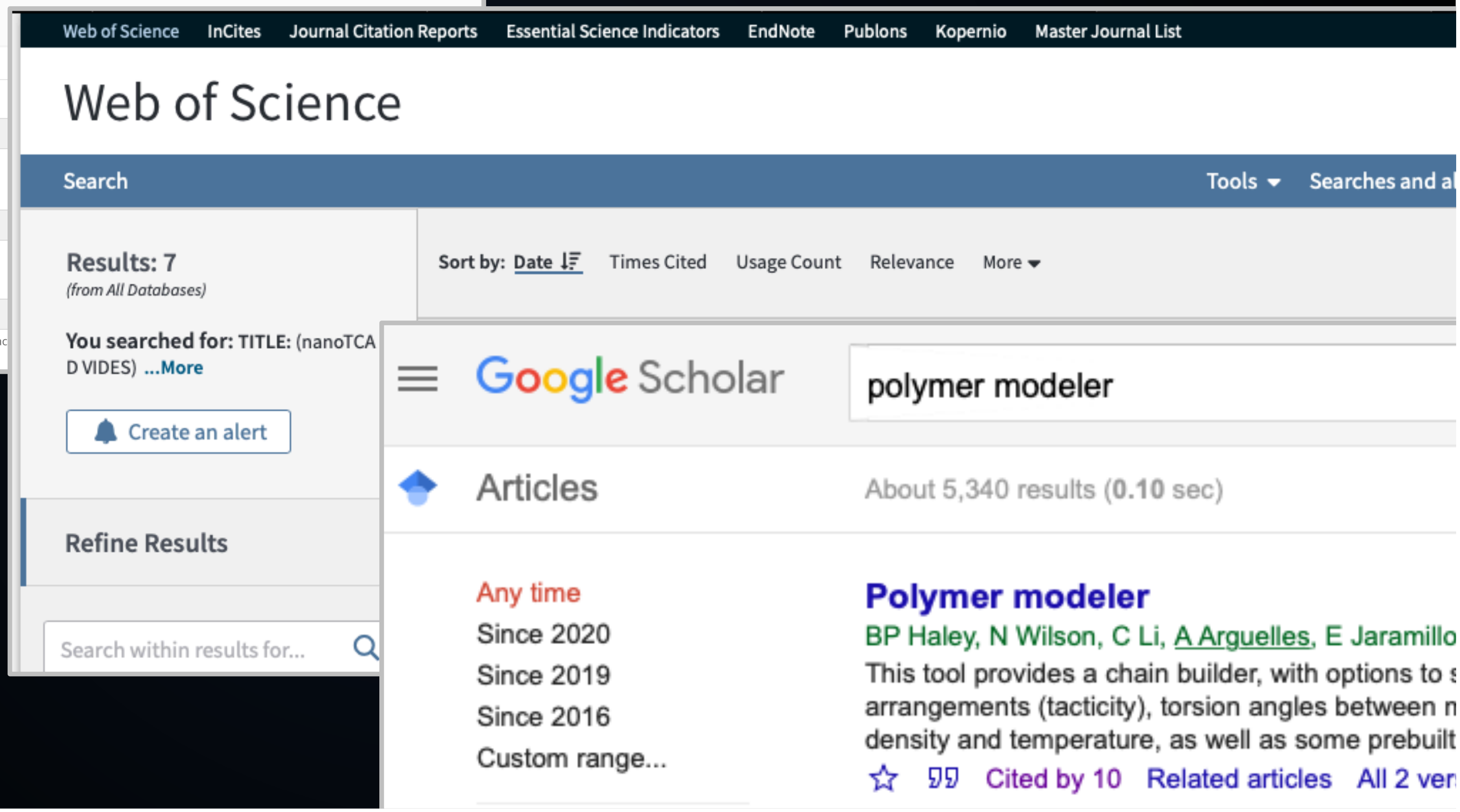
4. Post process results and plot



Publish for reproducibility and discoverability



Self-serve publication process:
https://nanohub.org/whypublish/



Tools & Apps indexed:

- Web of Science
- Google Scholar

Jupyter notebooks and data science

1. Query from Pymatgen

Pymatgen is an open-source library in python used for material analysis. Pymatgen is a powerful and popular resource that can be
Pymatgen makes querying these resources and obtain data from its internal database easy. We will start by querying the database

Making a query in Pymatgen requires the chemical symbol of the element, which are all listed in the cell above. From there, the pro
to learn more about the Element class.

In this example we will query the Young's modulus for the elements in the list "sample". You will be able to see the values with the
the "sample" elements.

```
In [2]: querable_pymatgen = ["atomic_mass", "poissons_ratio","atomic_radius", "electrical_resistivity","m
                             "brinell_hardness", "average_ionic_radius", "melting_point", "rigidity_modulu

        sample = ['Fe', 'Co', 'Ni', 'Cu', 'Zn']

        for item in sample:
            element_object = pymat.Element(item)
            print(item, element_object.youngs_modulus) # You can change "youngs_m

        #for item in sample:
        #    for i in querable_pymatgen:
        #        element_object = pymat.Element(item)
        #        print(item, i, getattr(element_object,i))

Fe 211.0 GPa
Co 209.0 GPa
Ni 200.0 GPa
Cu 130.0 GPa
Zn 108.0 GPa
```

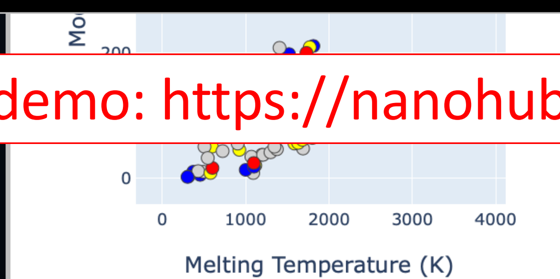
- **Exercise 1.** Modify the query above to extract Brinell hardness.
- **Exercise 2.** Uncomment the lines above to see all the properties of the

Remember: "Shift-Enter" to re-run the cell.

1. Obtain data by querying external databases

2. Filter, process, organize and visualize data

```
data = [trace0, legend_plot_FCC, legend_plot_BCC, legend_plot_HCP]

fig= go.Figure(data, layout=layout0)
```



Introduction to data science for materials science
- https://nanohub.org/tools/mseml

Data repositories
- https://nanohub.org/tools/matdatarepo

Dimensionality reduction
- https://nanohub.org/tools/dimredmatdecmp/

3. Train models (e.g. neural networks, random forests)

```
DEL
eural network will be initialized in a random manner, using a seed allows for reproducibility
zers.RandomNormal(seed=0)
l, the first layer must specify the input shape the model will expect;
ue is train_values.shape[1] which is the number
rties) and equals 17.

model.add(Dense(32, activation='relu', input_shape=(train_values.shape[1], ), kernel_initializer=kernel_init))
model.add(Dense(64, activation='relu', kernel_initializer=kernel_init))
#model.add(Dense(64, activation='relu', kernel_initializer=kernel_init))

opagation

ich metrics will evaluate the model's accuracy
model.compile(loss='mae', optimizer=optimizer, metrics=['mae'])
model.summary()

WARNING:tensorflow:From /apps/share64/debian7/anaconda/anaconda-6/lib/python3.7/site-packages/tensorflow/python/framewo
.python.framework.ops) is deprecated and will be removed in a future version.
Instructions for updating:
Colocations handled automatically by placer.
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
dense_1 (Dense)              (None, 32)                544
_________________________________________________________________
dense_2 (Dense)              (None, 64)                2112
_________________________________________________________________
dense_3 (Dense)              (None, 1)                 65
=================================================================
Total params: 2,721
Trainable params: 2,721
Non-trainable params: 0
_________________________________________________________________
```

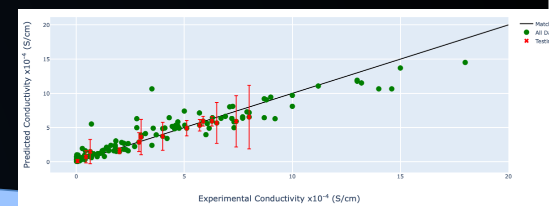
Hands-on demo: https://nanohub.org/tools/mseml

Data science for design of experiments

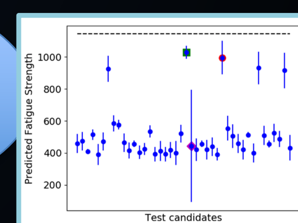
ML can help reduce the number of experiments to achieve a design goal

Existing data

Predictive model of state of knowledge



Information acquisition function



Query information source

Goal achieved?
YES: SUCCESS
NO: Add data & restart

Pandita P, Bilionis I, Panchal J. Journal of Mechanical Design. 2016 138, 111412.
Ling J, Hutchinson M, Antono E, Paradiso S, Meredig B. Int. Mat. and Manuf. Inn.. 2017 Sep 1;6(3):207-17.

Maximizing Li+ conductivity in solid oxides

Juan Carlos Verduzco

CITRINE INFORMATICS

nanoHUB

Materials Informatics resources fro

Authorship and credits

Based on work published on: Julia Ling et al.
Original notebooks from: *Logan Ward*, Argonne National Laborator
nanoHUB tools by: *Juan Carlos Verduzco* and *Alejandro Strachan*, N

Machine Learning Algorithms for Ionic Conductivity of LLZO-type Garnets

Authorship and credits

Based on work published on: Julia Ling et al.
Original notebooks from: *Logan Ward*, Argonne National Laboratory, and *Max Hutchinson*, Citrine Informatics. Their Github implementation can be found here.
nanoHUB tools by: *Juan Carlos Verduzco* and *Alejandro Strachan*, Materials Engineering, Purdue University
Database curated by: *Juan Carlos Verduzco*, Materials Engineering, Purdue University

Overview

In this notebook, I'll introduce different regression mach

the fa

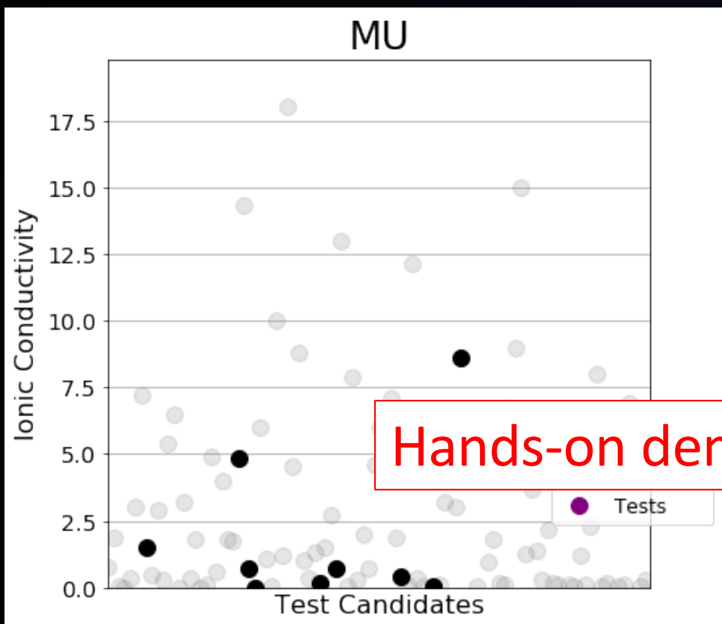
In [3]:
```
cdr = CitrineDataRetrieval('[CITRINE_API_KEY]') # Citrine Key

data = cdr.get_dataframe(criteria={'data_set_id': 184812}, print_properties_options=False) # LLZO Database
property_interest = 'Ionic Conductivity' # Property to be queried

display(data.head(n=10))
```

```
  0%|          | 0/199 [00:00<?, ?it/s]/apps/share64/debian7/anaconda/anaconda-6/lib/python3.7/site-packages/pandas/core/frame.py:
:

Sorting because non-concatenation axis is not aligned. A future version
of pandas will change to not sort by default.

To accept the future behavior, pass 'sort=False'.

To retain the current behavior and silence the warning, pass 'sort=True'.


100%|██████████| 199/199 [00:04<00:00, 46.71it/s]
```

	chemicalFormula	preparation	references	Crystallographic Structure	Ionic Conductivity	Ionic Conductivity-units
1	Li6.25La3Zr1.25Bi0.75O12	[{'name': 'SINTERING', 'details': [{'name': 'S...	[{'citation': 'http://arxiv.org/abs/1902.06831'}]	Cubic	2	10^-4 S/cm
2	Li6La2Nb1.5Y0.5O12	[{'name': 'SINTERING', 'details': [{'name': 'S...	[{'citation': '10.1021/jp304737x', 'doi': '10....	Cubic	1.85	10^-4 S/cm
3	Li6.25La3Zr2Ga0.25O12	[{'name': 'SINTERING', 'details': [{'name': 'S...	[{'citation': '10.1149/2.0171903jes', 'doi': '...	Cubic	14	10^-4 S/cm
4	Li6La3Ta1.5Y0.5O12	[{'name': 'SINTERING', 'details': [{'name': 'S...	[{'citation': '10.1039/C4CP00418C', 'doi': '10...	Cubic	1.26	10^-4 S/cm
5	Li6.55La3Zr2Ga0.15O12	[{'name': 'SINTERING', 'details': [{'name': 'S...	[{'citation': '10.1021/acsami.6b13902', 'doi':...	Mixed Cubic / Tetragonal	0.85	10^-4 S/cm
6	Li7.4La3Zr1.76W0.24O12	[{'name': 'SINTERING', 'details': [{'name': 'S...	[{'citation': '10.1016/j.electacta.2015.08.046...	Cubic	7.7	10^-4 S/cm
7	Li6.4La3Zr2Ga0.2O12	[{'name': 'SINTERING', 'details':	[{'citation':	Cubic	13.2	10^-4 S/cm

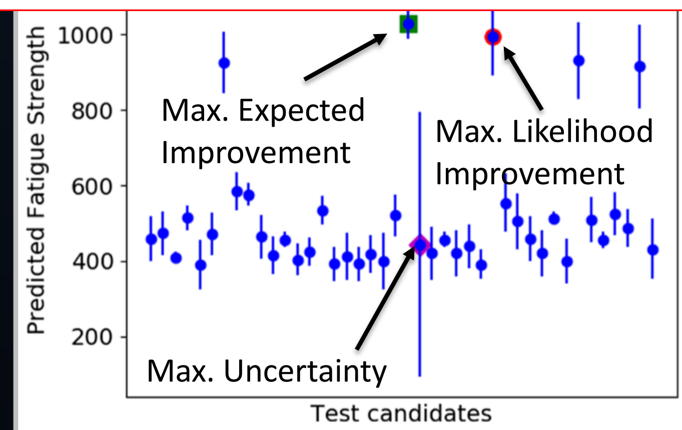
https://nanohub.org/tools/citrinetools

Find the best conductor with the fewest experiments



- Start with 10 data points (black) out of 100
- Train a ML model and evaluate on possible experiments
- Use acquisition functions to decide what experiment to reveal next
- Add result and iterate

Hands-on demo: https://nanohub.org/tools/citrinetools

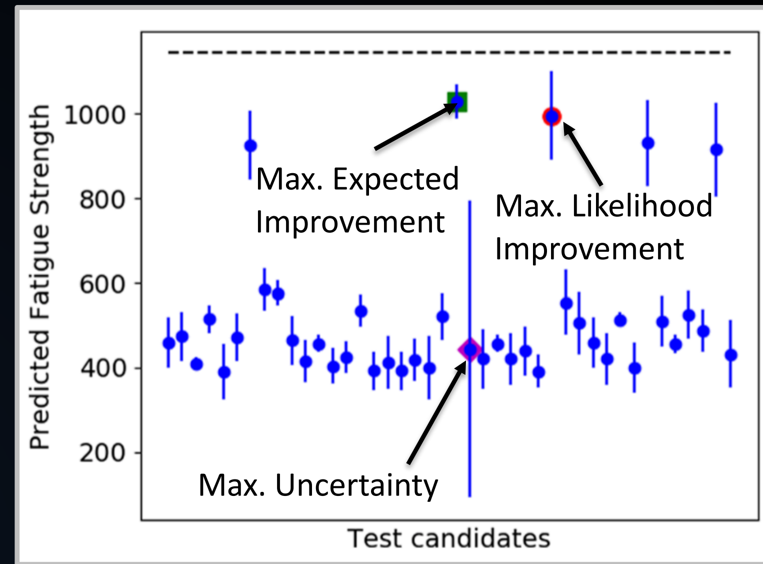


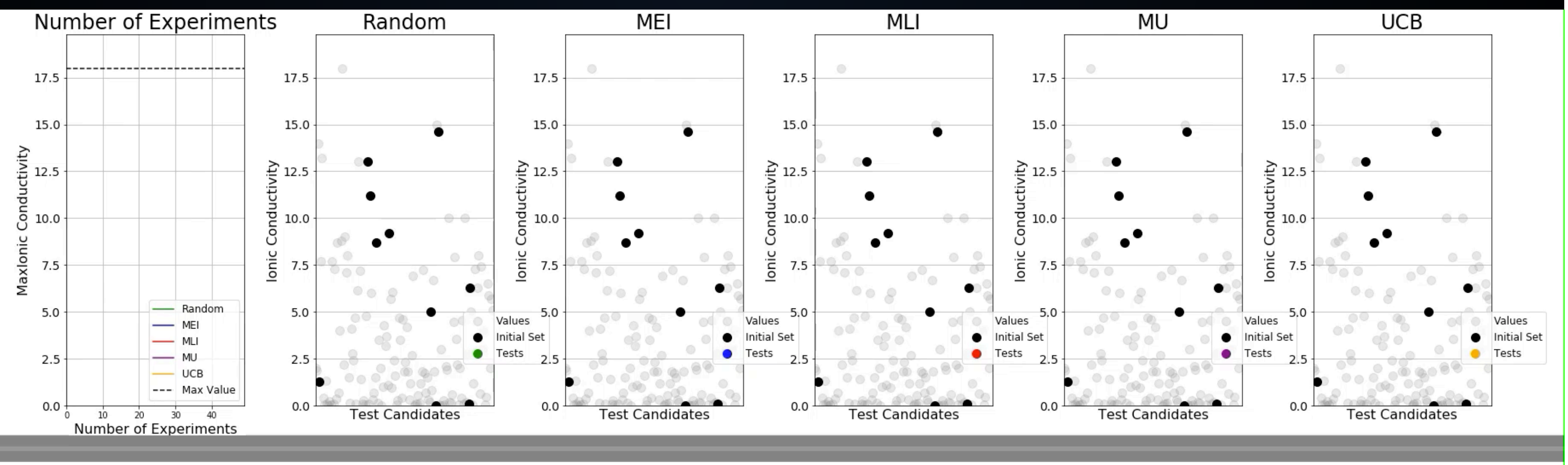
Ling J, Hutchinson M, Antono E, Paradiso S, Meredig B. Int. Mat. and Manuf. Inn.. 2017 Sep 1;6(3):207-17.

Finding best conductor with the fewest experiments

- Run the notebook online
- Adapt it to your specific problem

https://nanohub.org/tools/citrinetools

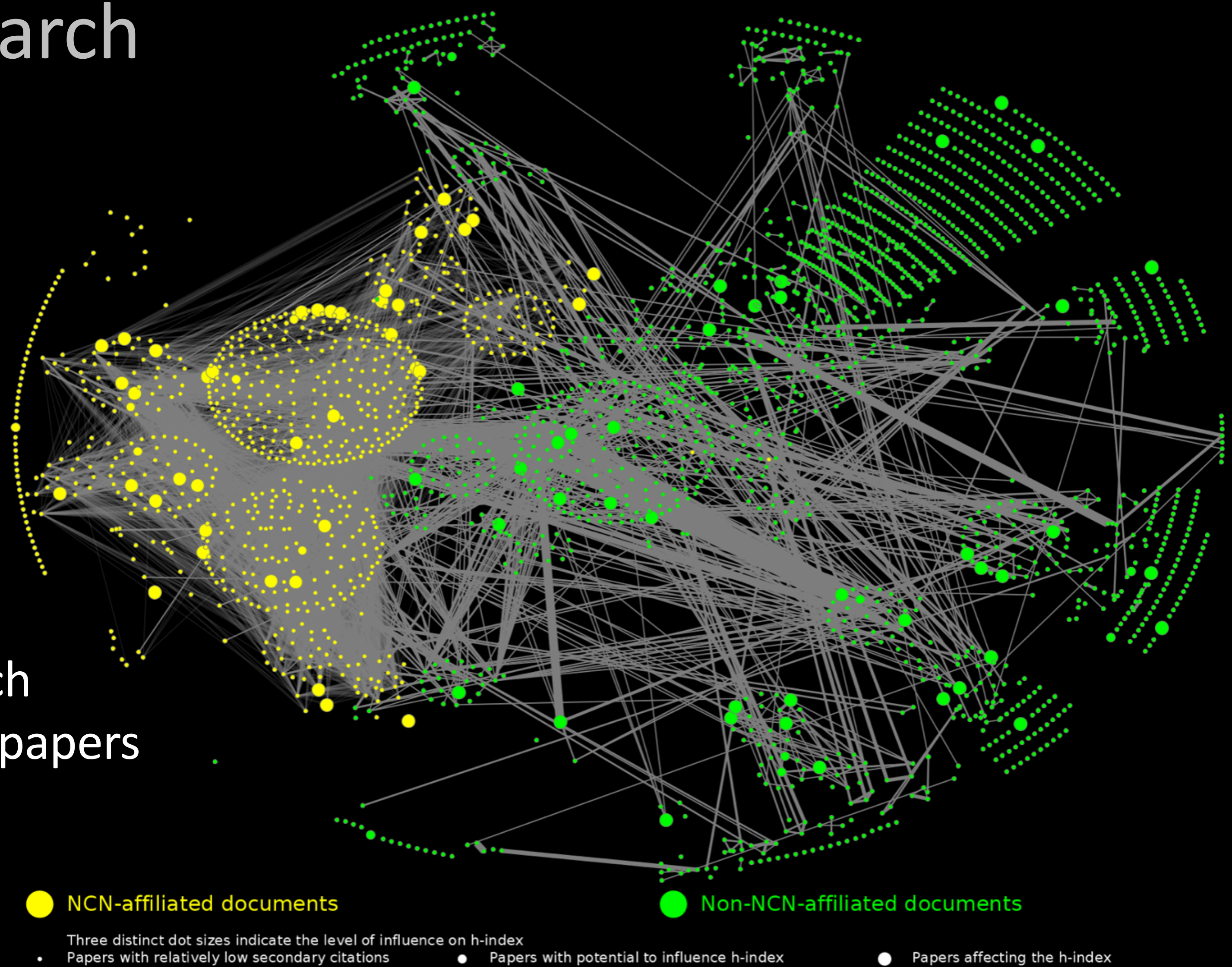


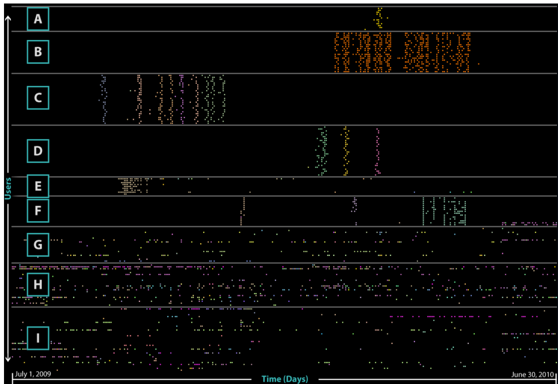


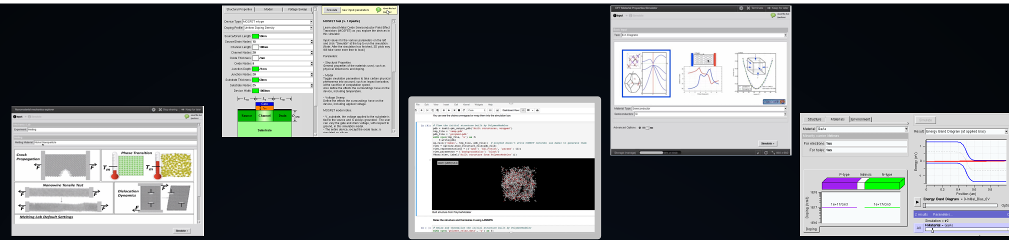
Impact on research

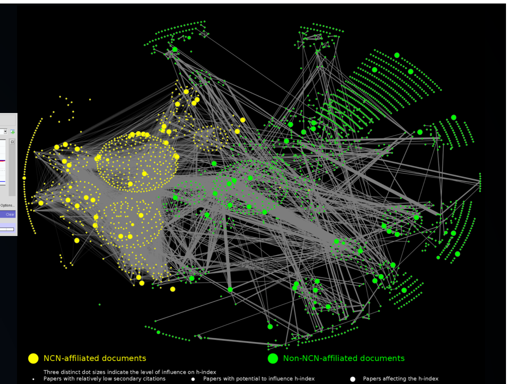
2,400+ citations
5,900+ authors

2,100+ use in research
1,600+ Outside NCN papers





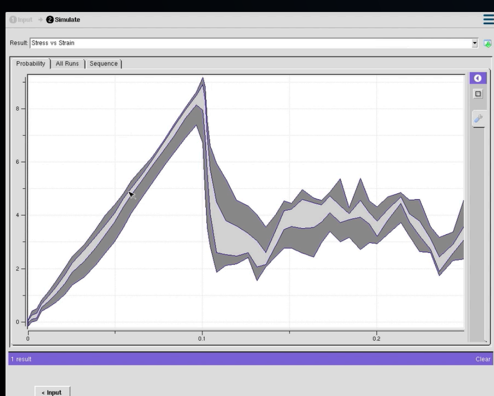




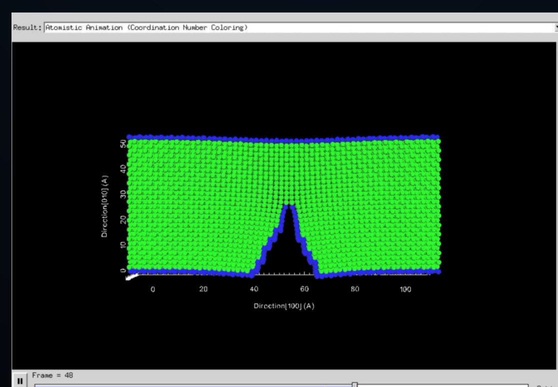
Recent innovations

making Simulations and Data More Powerful

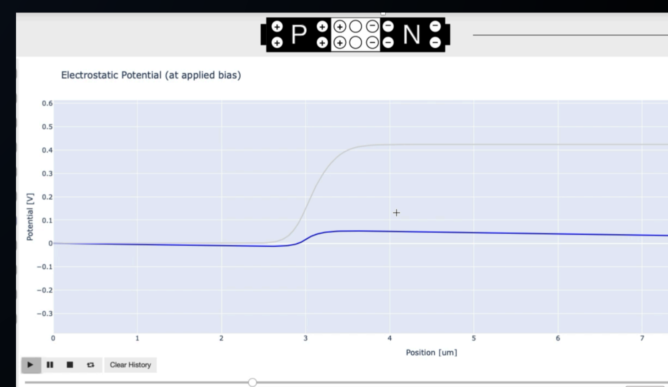
Automatic uncertainty quantification



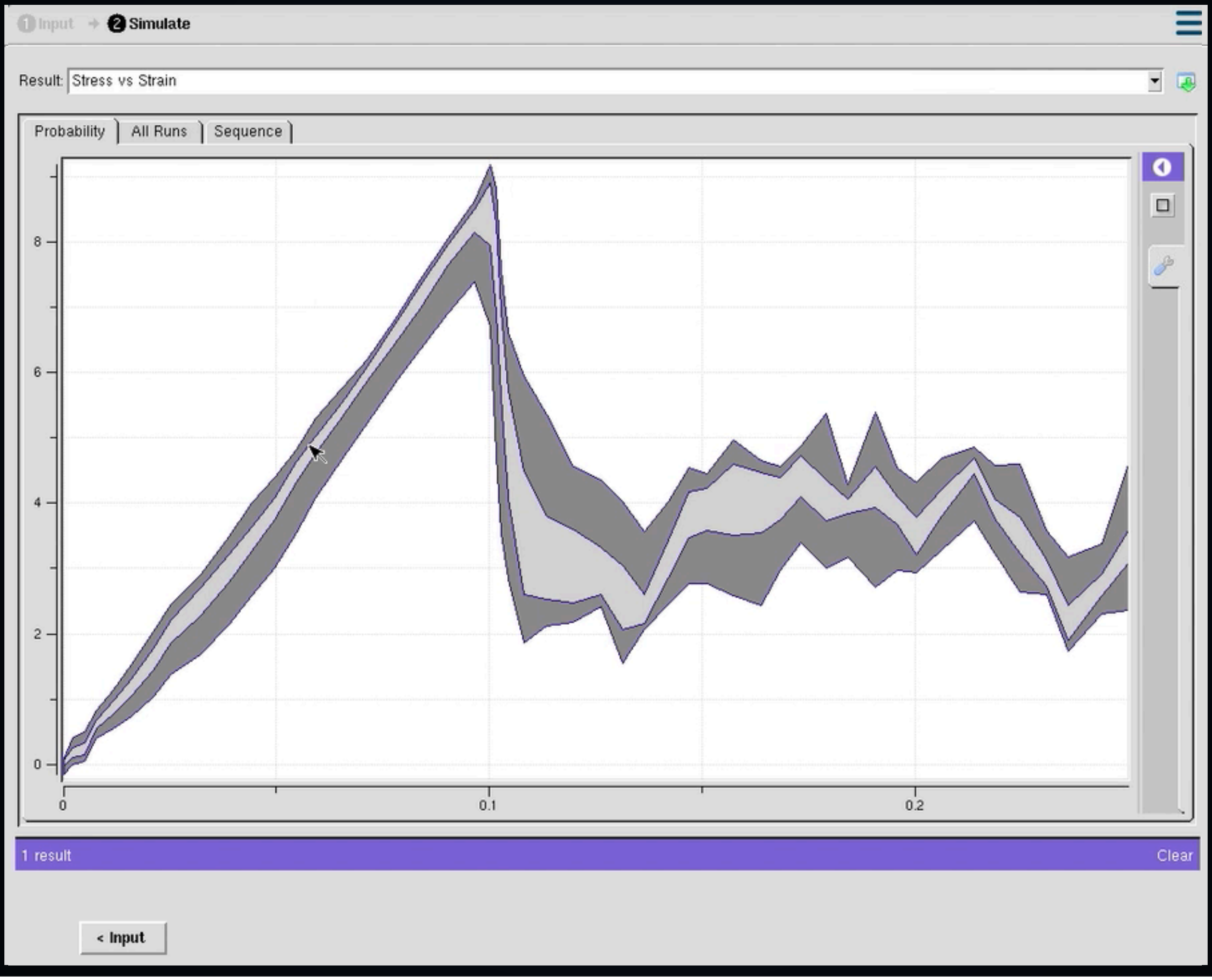
Instant feedback via simulation cache & simulation result database



New ways of delivering simulation & data

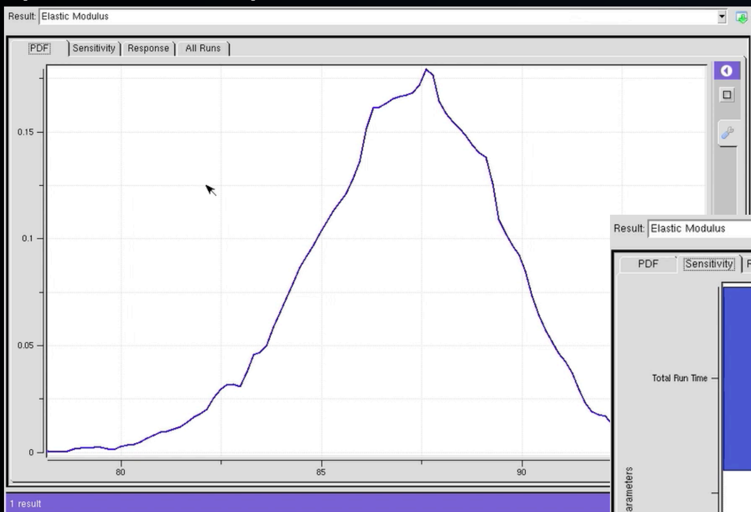


Automatic uncertainty quantification

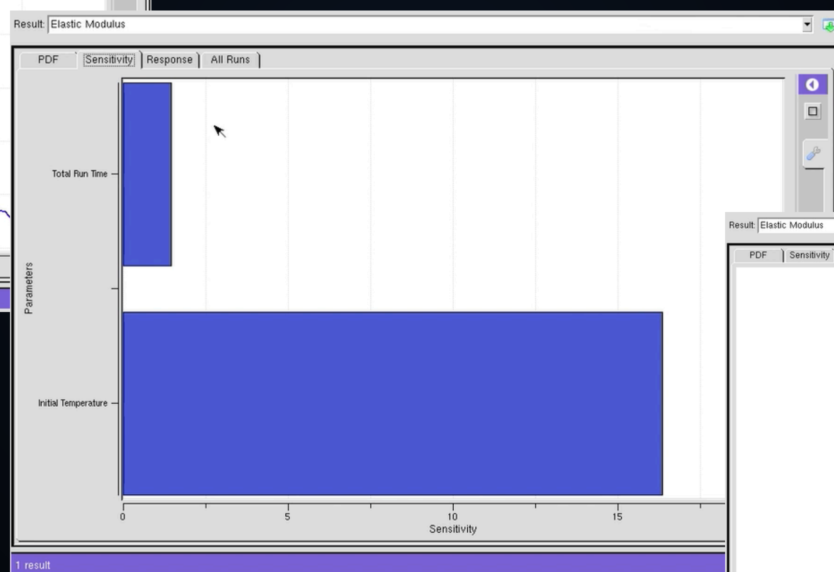


Automatic uncertainty quantification

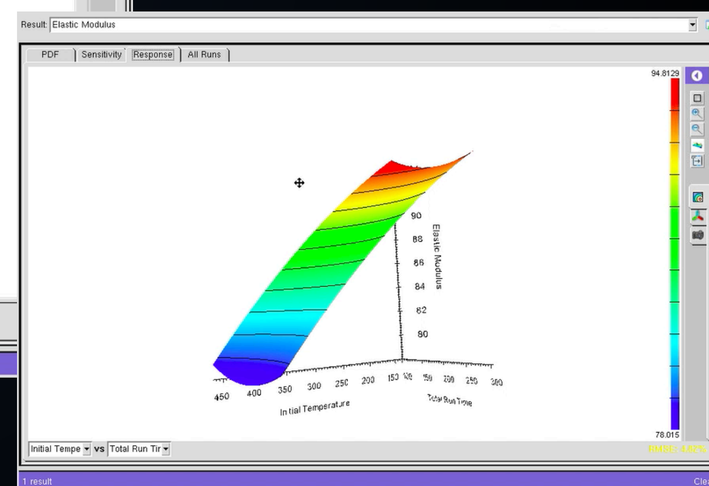
Uncertainties in predicted quantities



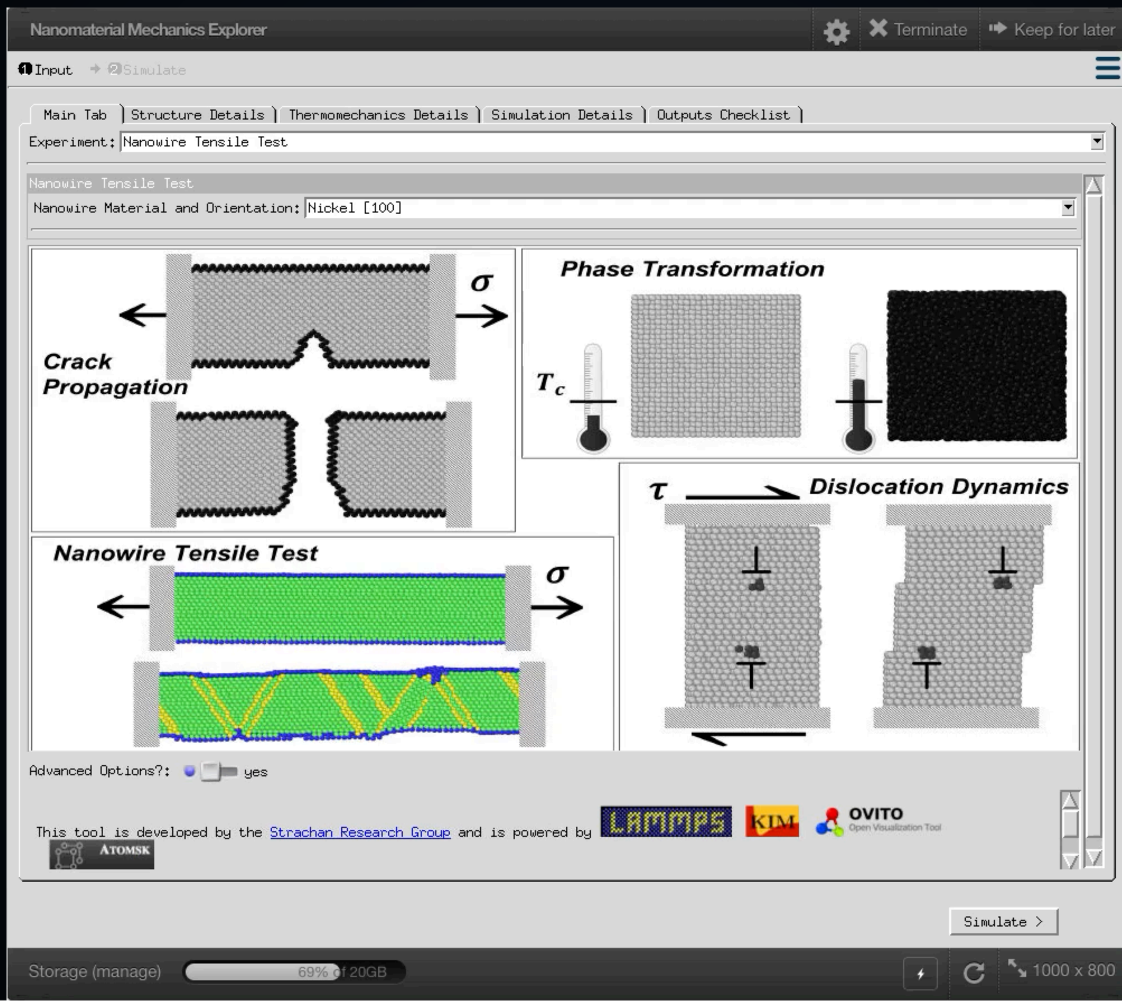
Sensitivity analysis



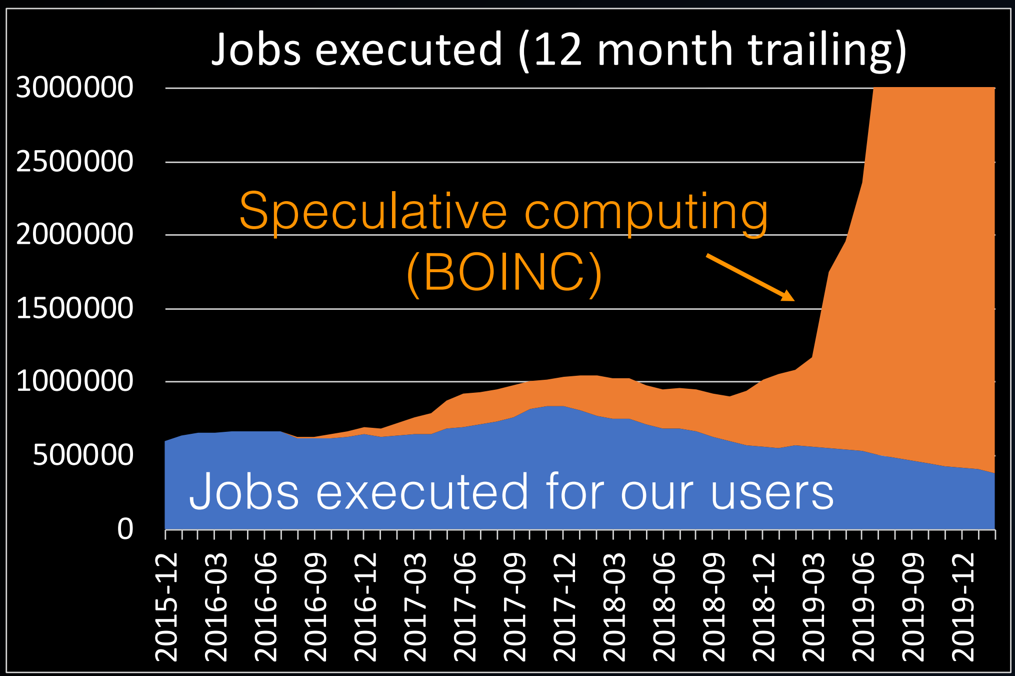
Surrogate models



Instant feedback via simulation caching



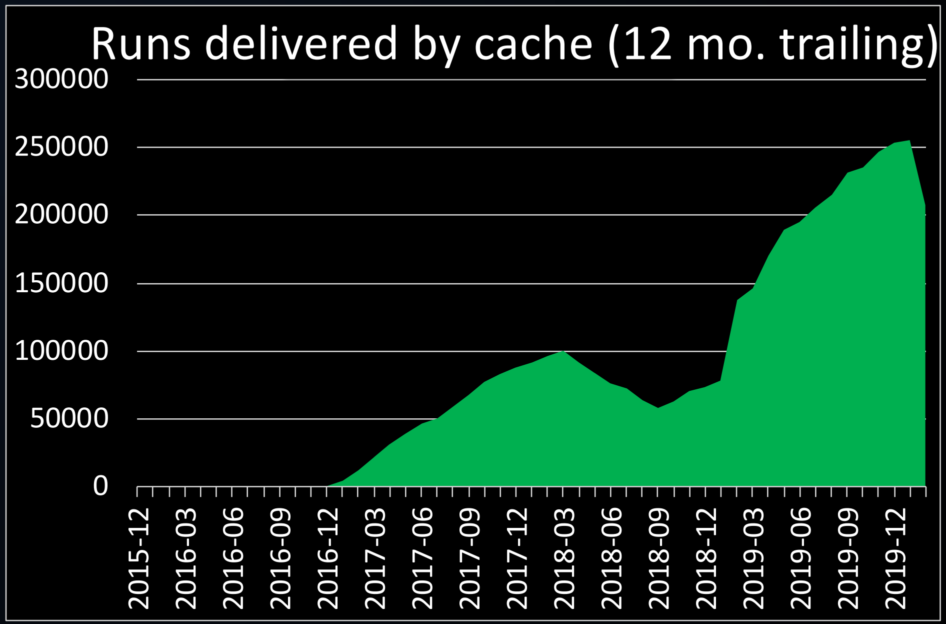
Simulation caching & computational resources



Are we serving fewer simulations?

No!

nanoHUB is going green!



Join your colleagues in nanoHUB & learn more

Materials science: https://nanohub.org/groups/materials

Welcome to the Materials Science group!

Here, you can find many resources for materials science research and education, including complete courses, seminars and tutorials on specialized topics, simulation tools that run in the cloud using nanoHUB's computing resources, and learning activities that use simulations.

Be sure to catch the Summer 2020 MSE Education Webinar Series, July through August!

There are also activities for Materials Science outreach that you can use to teach high school and middle school students about the wonders of materials science.

Everyone is welcome to explore the nanoHUB resources collected here.

Materials science faculty and researchers are invited to join this community to take part in discussions and access additional functionalities of this nanoHUB group.

Scroll down this page or use the menu to find links to the different materials science resources.

Data science & machine learning: https://nanohub.org/groups/ml

Hands-on Data Science and Machine Learning Training

By Alejandro Strachan[1], Saaketh Desai[2]

1. Materials Engineering, Purdue University, West Lafayette, IN 2. Purdue University, West Lafayette, IN

Edit

View Series

Video podcast

Slides/Notes podcast

https://nanohub.org/resources/33245

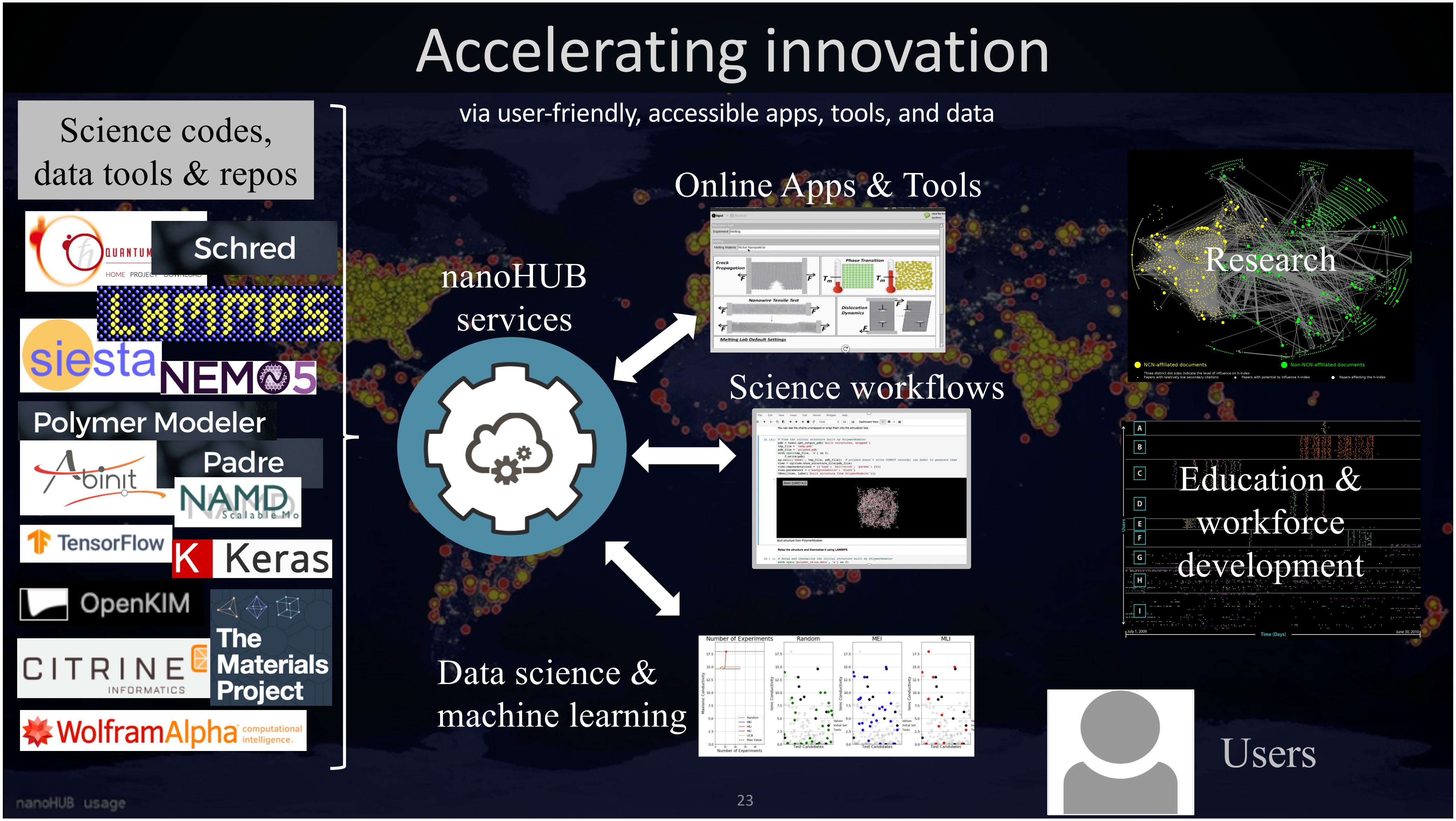
Simulation Tools | Educational Material and Training | Research Seminars | Machine Learning Workshop

Hands-on Machine Learning Workshop

We conducted an online, hands-on machine learning workshop in April 2020. Click here to find out more.

Educational Material and Training

Overview: Machine Learning for Materials Science


Accelerating innovation
via user-friendly, accessible apps, tools, and data
Science codes, data tools & repos
Quantum
Schred
LAMMPS
siesta
NEMO5
Polymer Modeler
Abinit
Padre
NAMD
TensorFlow
Keras
OpenKIM
The Materials Project
CITRINE INFORMATICS
WolframAlpha computational intelligence.
nanoHUB services
Online Apps & Tools
Science workflows
Data science & machine learning
Research
Education & workforce development
Users
nanoHUB usage
23