Get your feet wet!

Introduction to the nanoHUB Simulation Environment

Tanya Faltens, Ph.D.
Network for Computational Nanotechnology
Purdue University

Overview

- What is nanoHUB
- Simulation Examples
  - Carbon nanostructures
  - Molecular vibrations
  - Biological molecules
- Jupyter notebooks
- The nanoHUB dashboard

What is nanoHUB?

An open-access cyberinfrastructure

National Science Foundation
EEC 1227110
Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation

A library of STEM resources

A Simulation Toolbox

A Workspace

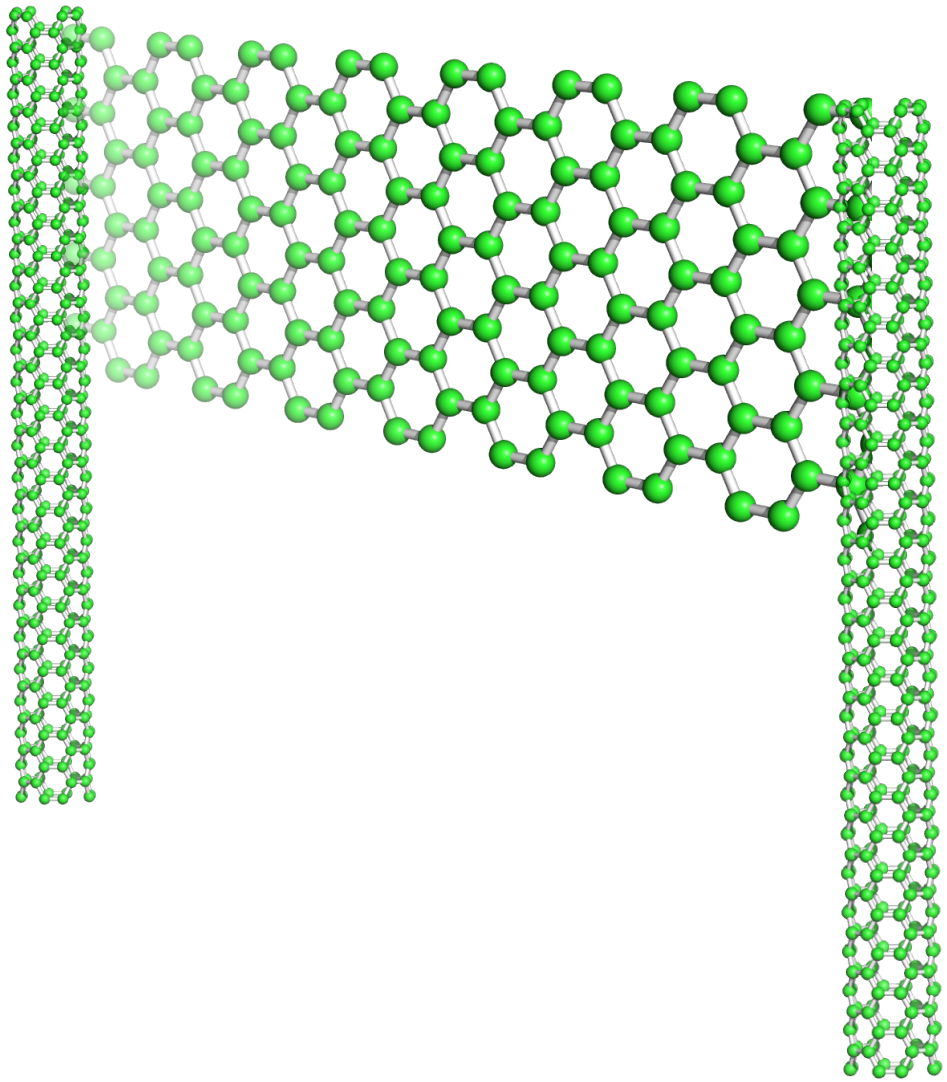
A Publishing Platform


A Scientific Community
1.9 M annual users worldwide
HUB usage 2016-01-01 00:00:00

QR Code for the Handout

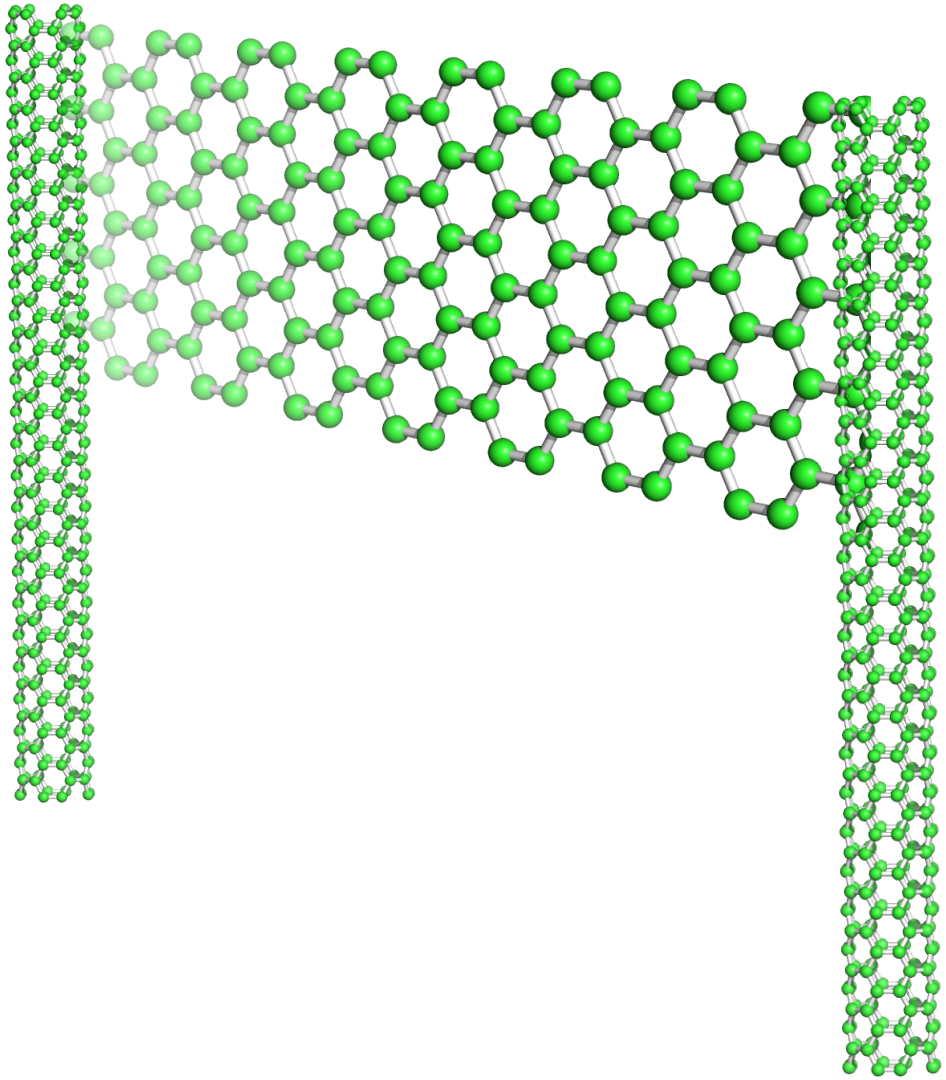
https://tinyurl.com/nH2021Mar16

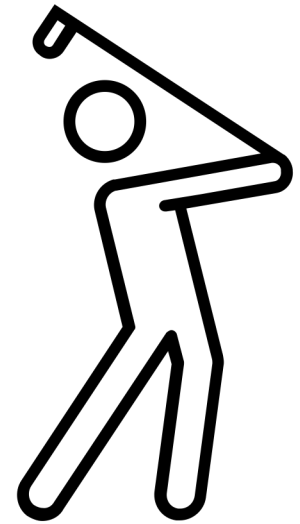
Fun with Carbon Nanostructures



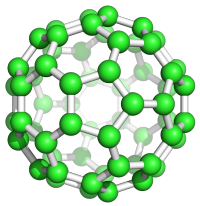
Fun with Carbon Nanostructures

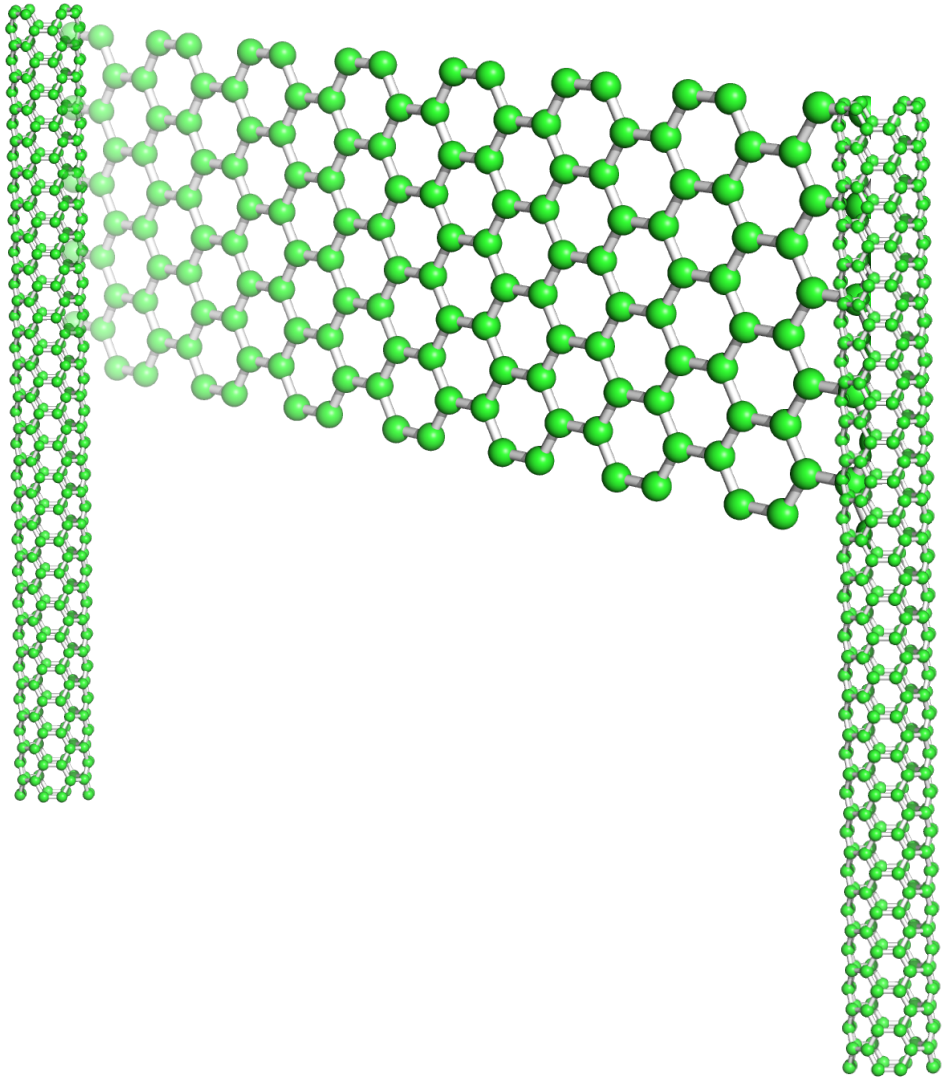
choose a player

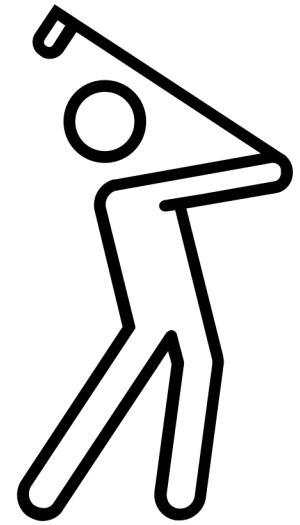




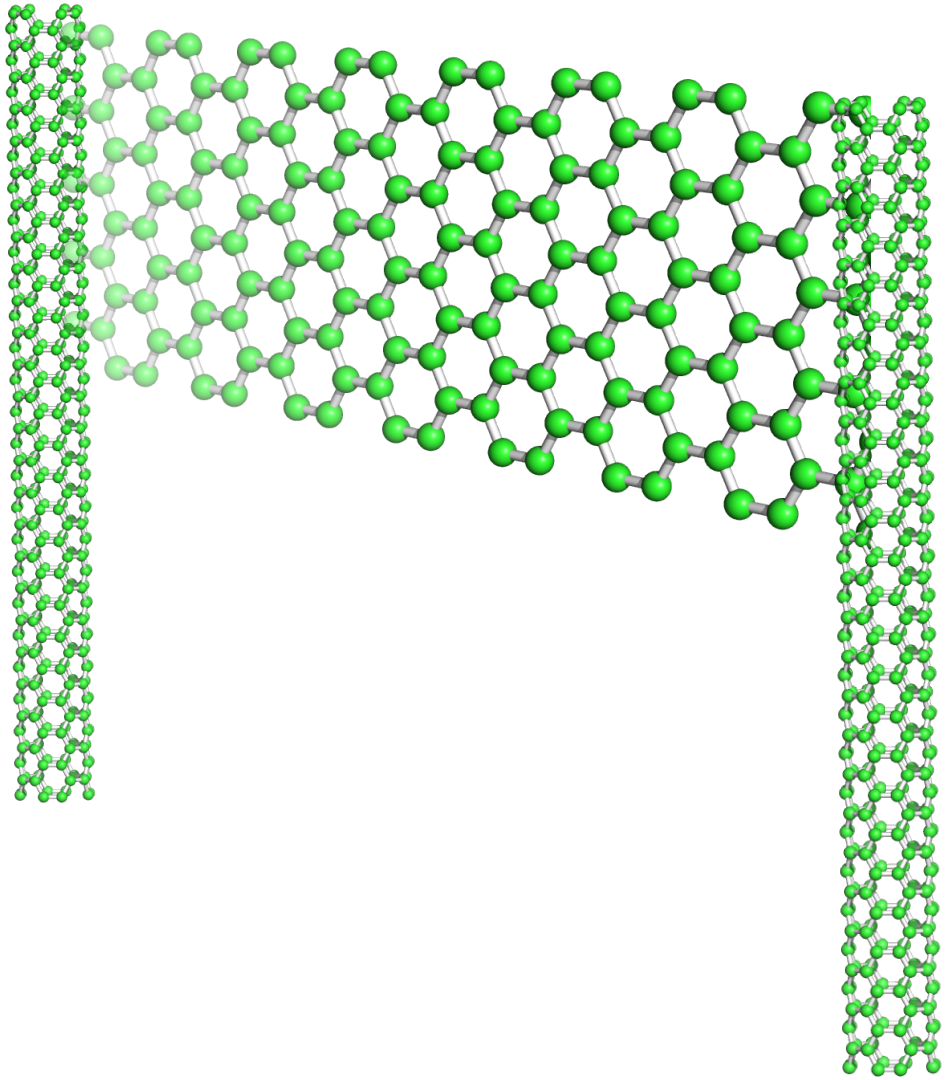
Fun with Carbon Nanostructures

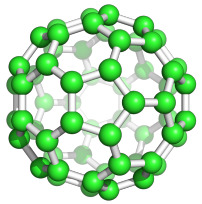




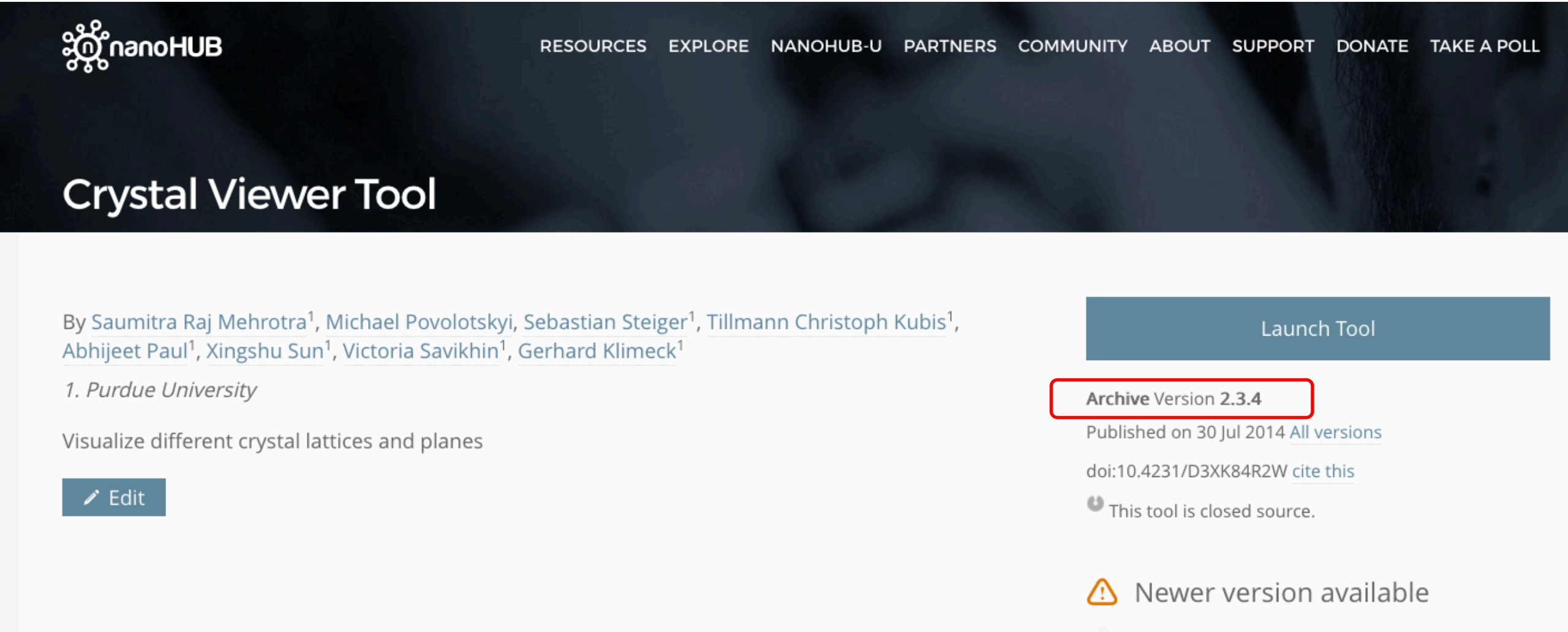


Fun with Carbon Nanostructures

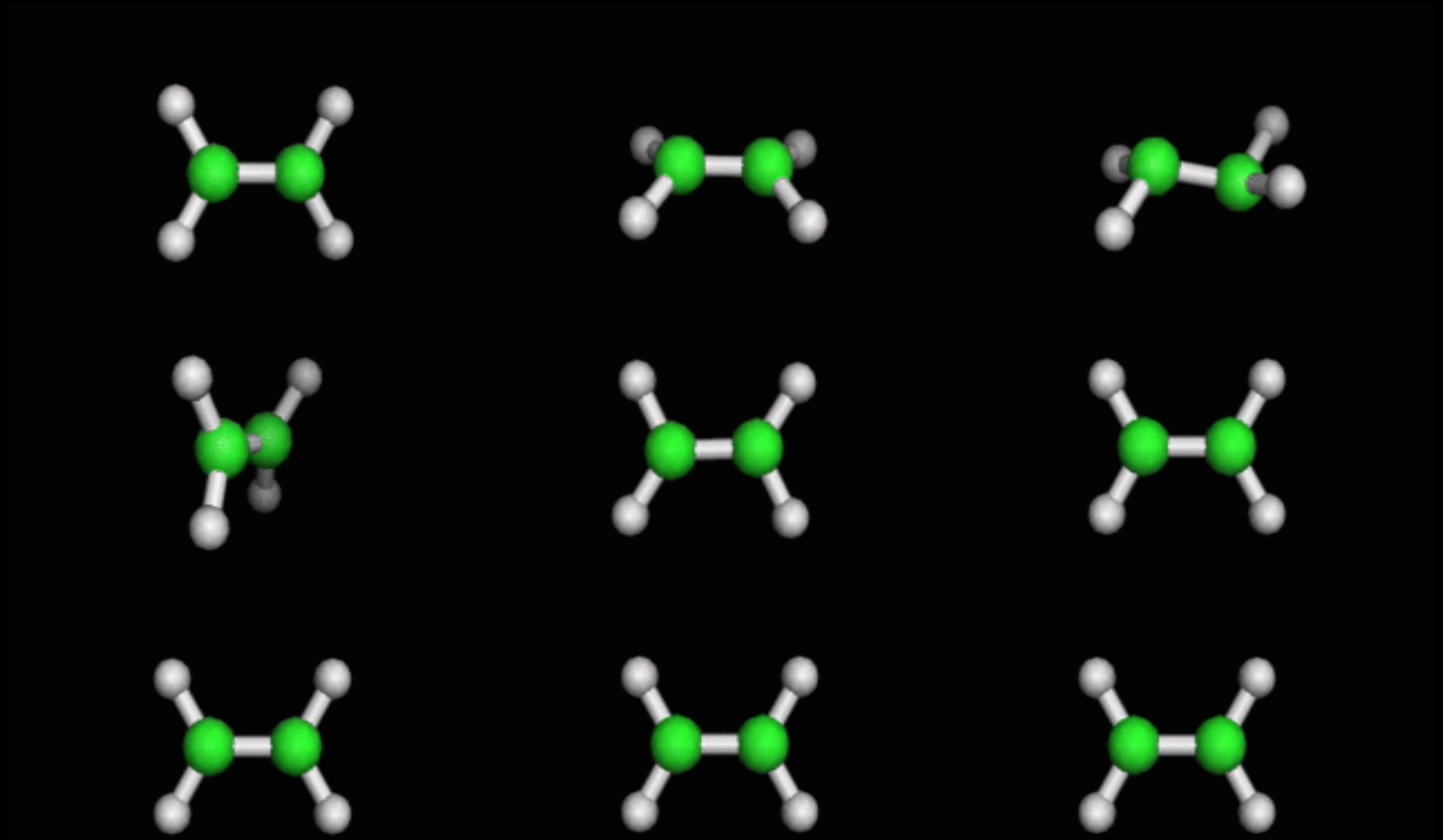




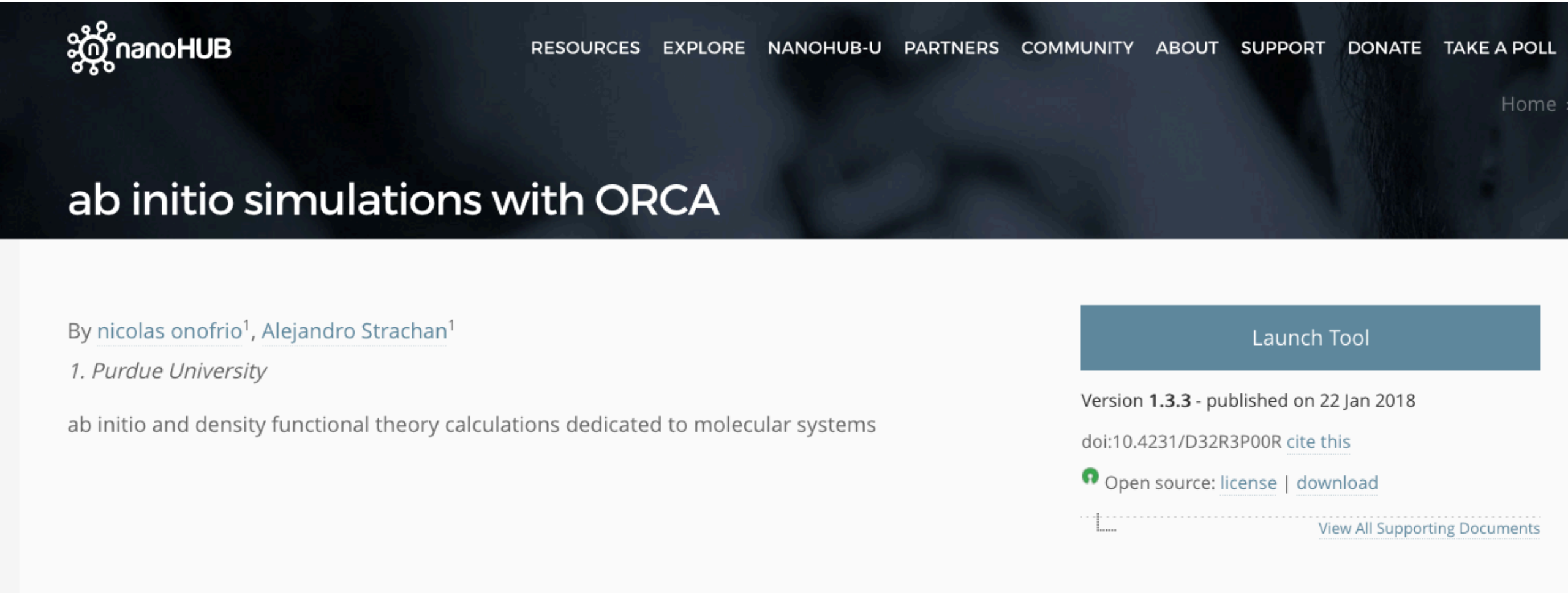
Crystal Viewer 2.3.4



Ethene Vibrational Modes

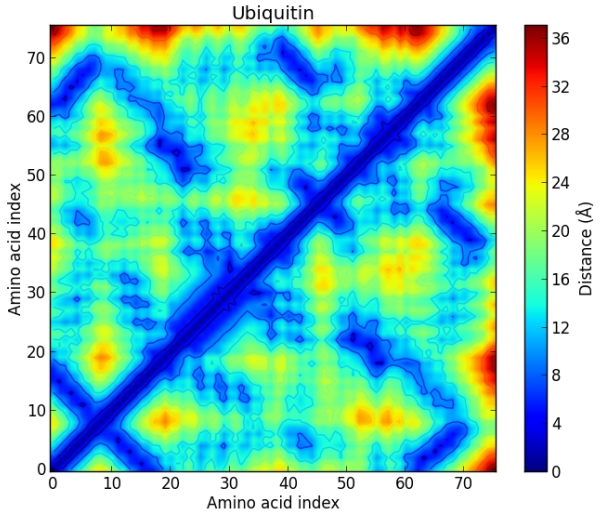


ab initio simulations with ORCA

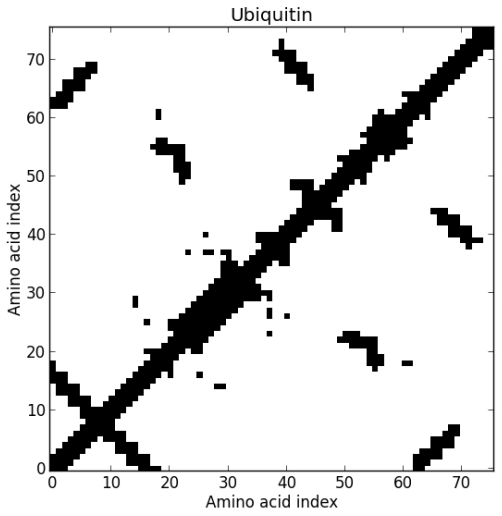


Ubiquitin (1UBQ)

Molecular Structure

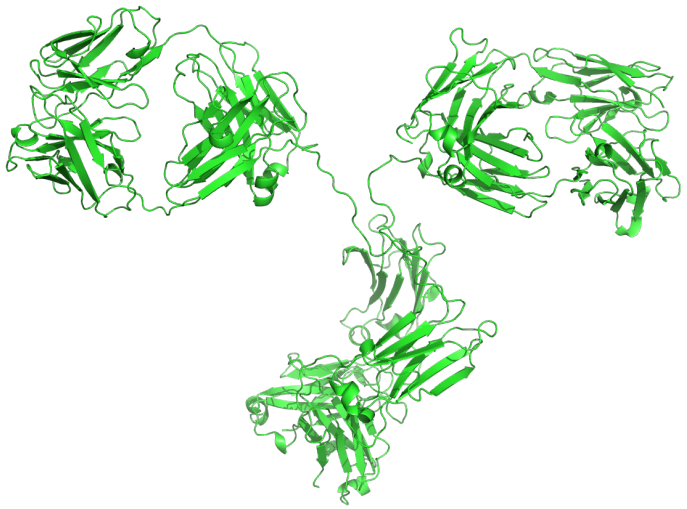


Distance Map

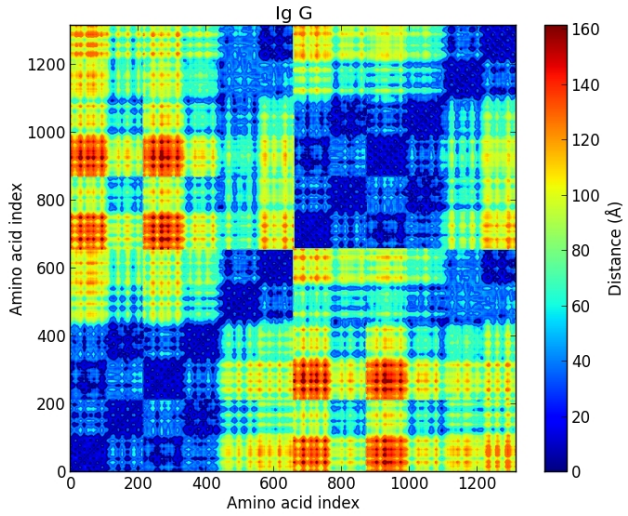


Contact Map

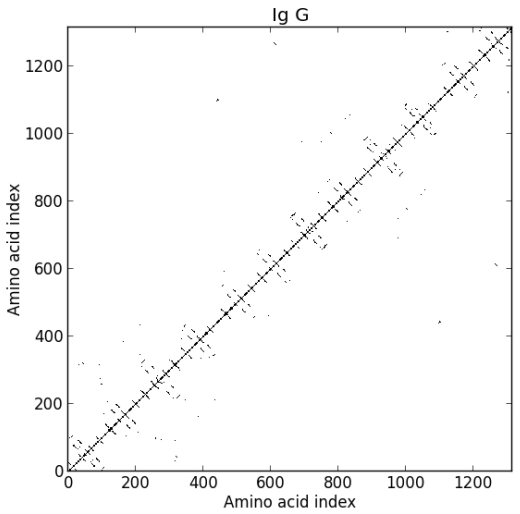
Immunoglobulin G (1IGT)



Molecular Structure

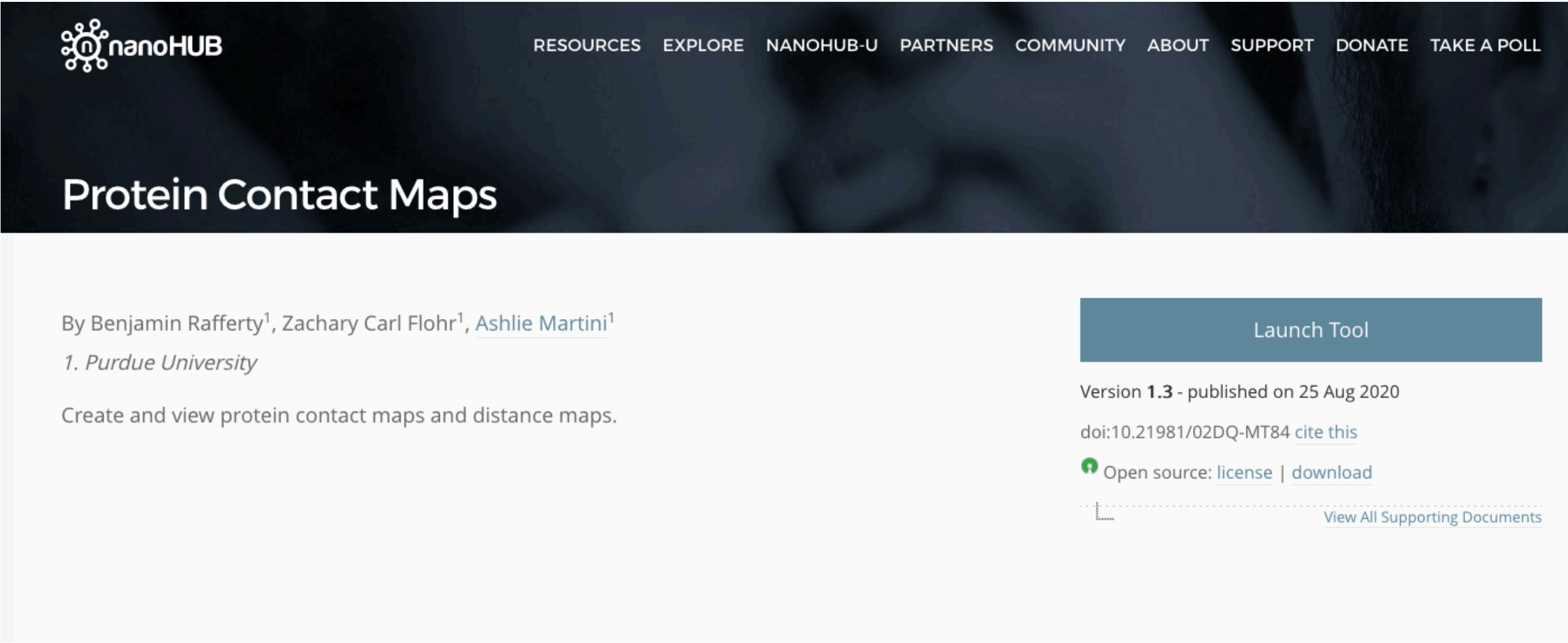


Distance Map



Contact Map

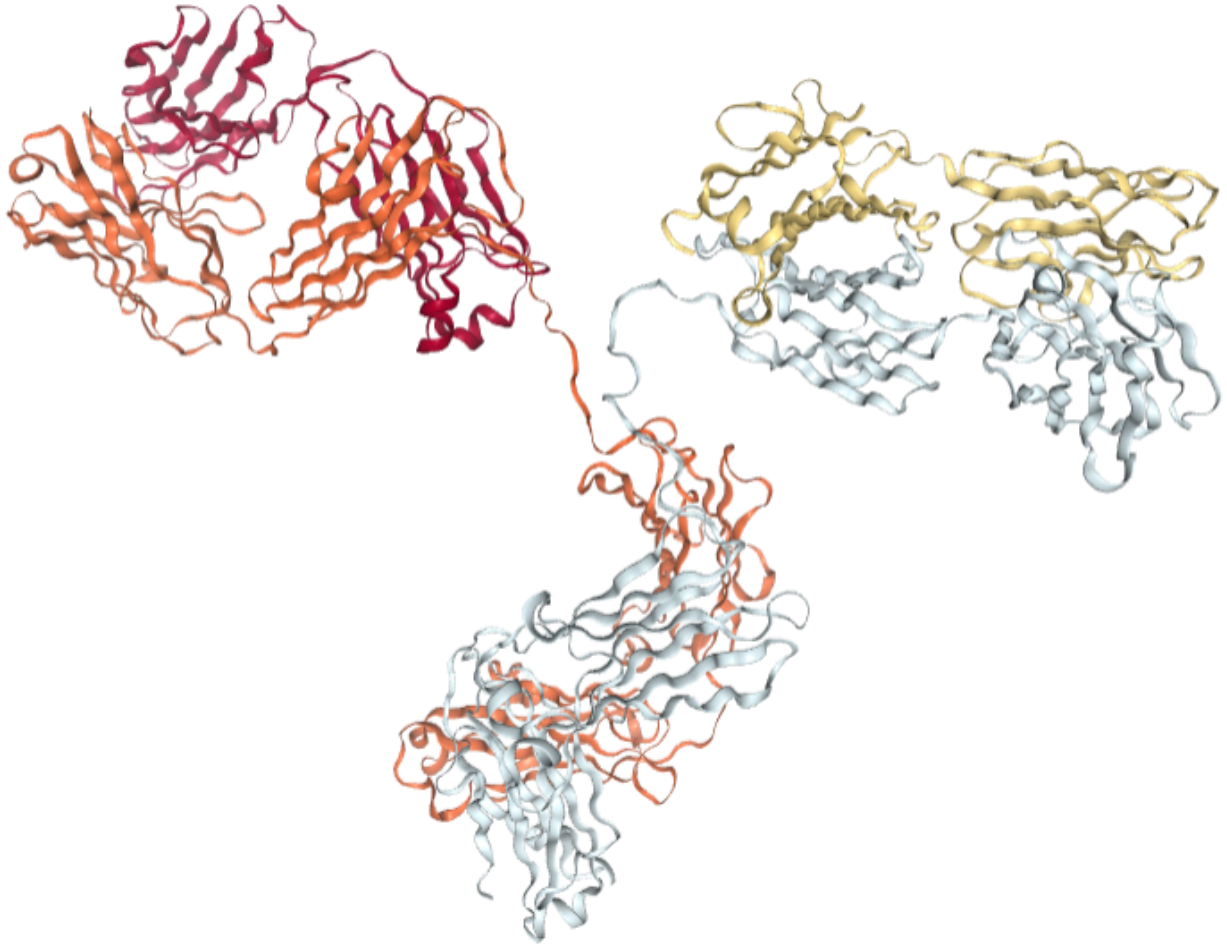
Protein Contact Maps



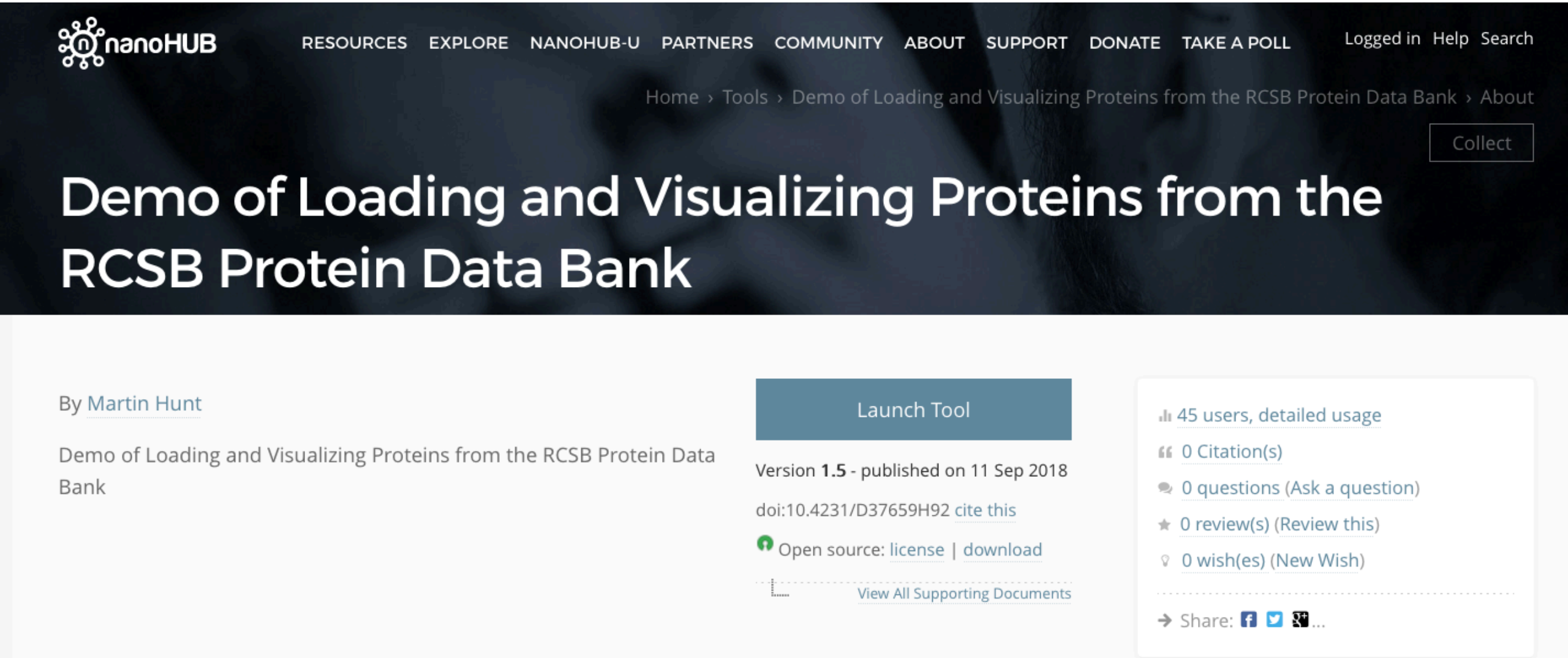
RCSB PDB
PROTEIN DATA BANK

A demonstration of loading a protein from the RCSB Protein Data Bank and visualizing it in a Jupyter Notebook. Click on the protein to enter fullscreen mode.

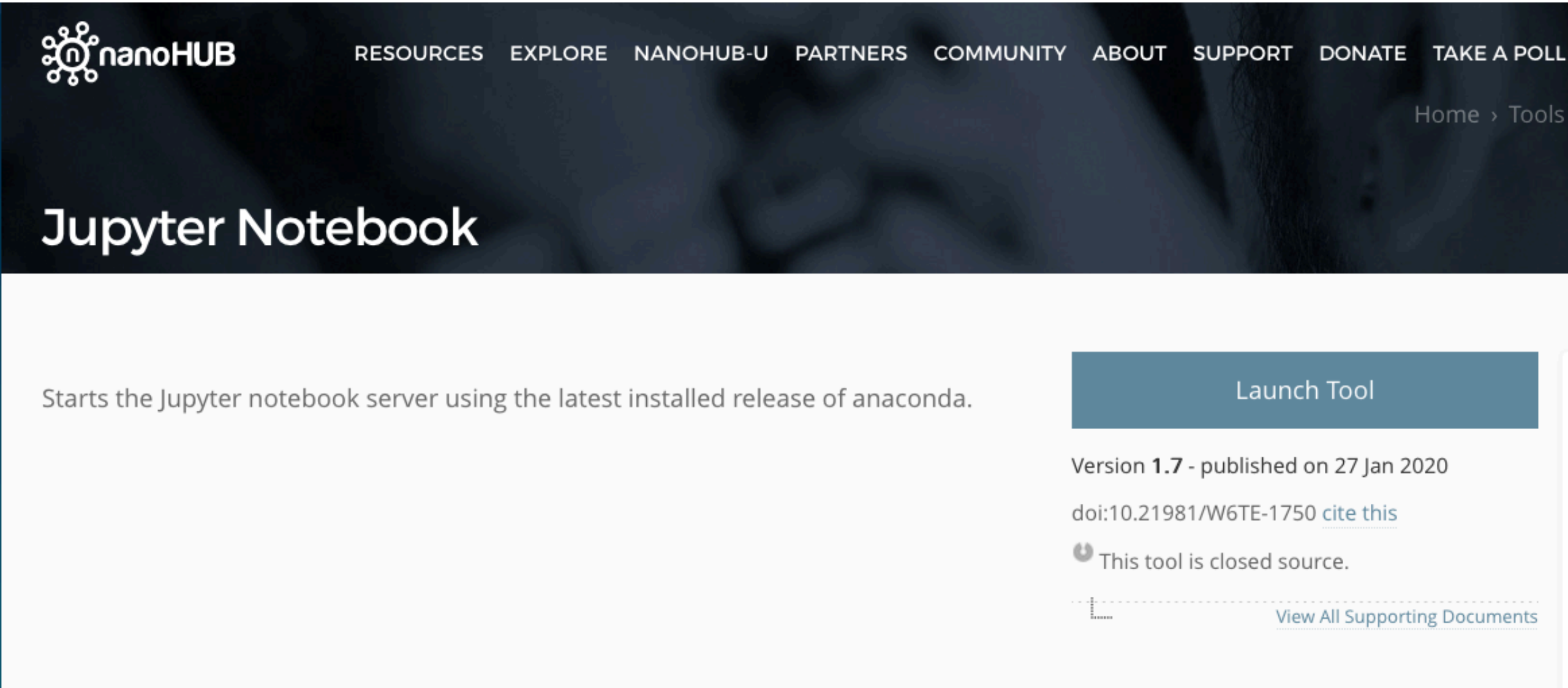
Immunoglobulin G (1IGT)



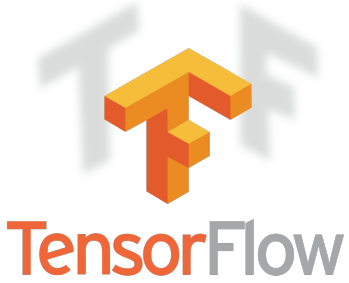
Demo of Loading and Visualizing Proteins from the RCSB Protein Data Bank



Jupyter notebooks in nanoHUB



Jupyter Kernels, Libraries, & Packages in nanoHUB include…



nanoHUB Provides Connections to Data

COD

MATERIALS DATA FACILITY

RCSB PDB
PROTEIN DATA BANK

WolframAlpha computational intelligence

KIM

CITRINE INFORMATICS

The Materials Project

Managing your Jupyter notebooks

See the handout:

Setting up Your nanoHUB File Structure in Jupyter Notebooks

biopython Tutorial

2.2 Working with sequences

Disputably (of course!), the central object in bioinformatics is the sequence. Thus, we'll start with a quick introduction to the Biopython mechanisms for dealing with sequences, the `Seq` object, which we'll discuss in more detail in Chapter 3.

Most of the time when we think about sequences we have in my mind a string of letters like '`AGTACACTGGT`'. You can create such `Seq` object with this sequence as follows - the ">>>" represents the Python prompt followed by what you would type in:

```
>>> from Bio.Seq import Seq
>>> my_seq = Seq("AGTACACTGGT")
>>> my_seq
Seq('AGTACACTGGT')
>>> print(my_seq)
AGTACACTGGT
```

The `Seq` object differs from the Python string in the methods it supports. You can't do this with a plain string:

```
>>> my_seq
Seq('AGTACACTGGT')
>>> my_seq.complement()
Seq('TCATGTGACCA')
>>> my_seq.reverse_complement()
Seq('ACCAGTGTACT')
```

mendeleev

Mendeleev Periodic Table Tutorial

Plotting tutorial

The plotting module is based on the Bokeh package and enables visualization of various properties of elements. The plotting in the form of periodic table is done by the `periodic_plot` function that takes a pandas `DataFrame`.

To embed the plots in the jupyter notebook first the `BokehJS` needs to be started.

[1]:
```
from bokeh.plotting import output_notebook, output_file
```

[2]:
```
output_notebook()
```

BokehJS 2.0.1 successfully loaded.

Now we can get the data from `mendeleev` as a pandas `DataFrame` through `get_table` method

[3]:
```
from mendeleev import get_table
from mendeleev.plotting import periodic_plot
```

[4]:
```
ptable = get_table('elements')
```

To plot the default table pass the `ptable` to the `periodic_plot` function

[5]:
```
periodic_plot(ptable)
```

Machine Learning for Materials Science

Collect

Machine Learning for Materials Science: Part 1

By Juan Carlos Verduzco Gastelum[1], Alejandro Strachan[1], Saaketh Desai[1]

1. Purdue University

Machine learning and data science tools applied to materials science

Launch Tool

Version **1.3** - published on 01 Apr 2020

doi:10.21981/WGQC-3249 cite this

Open source: license | download

View All Supporting Documents

2131 users, detailed usage

0 Citation(s)

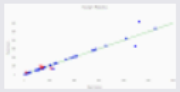
1 question (Ask a question)

2 review(s) (Review this)

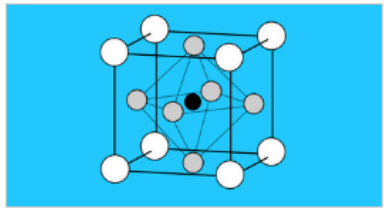
0 wish(es) (New Wish)

Share: ...

About | Usage | Citations | Questions | Reviews | Wishlist | Versions | Supporting Docs



Materials Science



Category

Tools

Published on

01 Apr 2020

Introduction to Machine Learning for Materials Science

The tutorials here will give you an insight into the usage of Machine Learning to approach problems related to materials science.

- **Get started** Click on the links below to begin each tutorial.
- **Important** To exit individual tutorials and return to this page, use File -> Close and Halt. "Terminate Session" (top right) will kill your entire Jupyter session.

Querying databases, Organizing and Plotting Data:

- Query Pymatgen and Mendeleev for properties like Young's modulus and melting temperature
- Organize data into Pandas dataframes and python dictionaries and plot using Plotly

Linear Regression to predict material properties:

- Perform linear regression using the scikit learn package and predict Young's modulus
- Visualize trends in data and 'goodness of fit' of linear model

Neural Network Regression to predict material properties:

- Use neural networks to perform non-linear, higher order regression
- Visualize trends and compare non-linear model to linear regression

Neural Network Classification to predict crystal structures:

- Use neural networks to classify elements according to their crystal structures

Hands-on Data Science and Machine Learning Training

https://nanohub.org/resources/33245

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

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

View Series

Video podcast

Slides/Notes podcast

- Introduction to Jupyter Notebooks, Data Organization and Plotting
- Repositories and Data Management
- Hands-on Supervised Learning: Part 1 - Linear Regression and Neural Networks
- Hands-on Supervised Learning: Part 2 - Classification and Random Forests
- Hands-on Sequential Learning and Design of Experiments

Saaketh Desai

Zachary McClure

Michael Sakano

Juan Carlos Verduzco

A course that uses Jupyter notebook tools

Slide content from Alejandro Strachan

Video tutorials to accompany the mseml notebooks