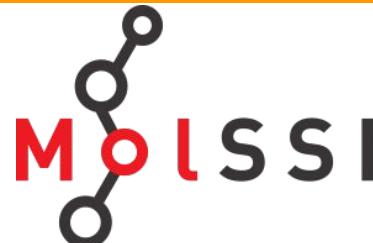


# Moving from User to Developer: Analyzing Molecular Simulations and Building New Tools



MD  
**ANALYSIS**

and



**MolSSI**

## Hybrid Workshop

(Arizona State University)

24-25 June 2024

**NUMFOCUS**  
[FISCALLY SPONSORED PROJECT]

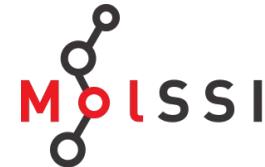
Chan  
Zuckerberg  
Initiative

**ASU** Center for  
Biological Physics  
Arizona State University

# Workshop Team



- Irfan Alibay
- Oliver Beckstein
- Mike Henry\*
- Ian Kenney
- Fiona Naughton
- Jenna Swarthout Goddard\*
- Yuxuan Zhuang



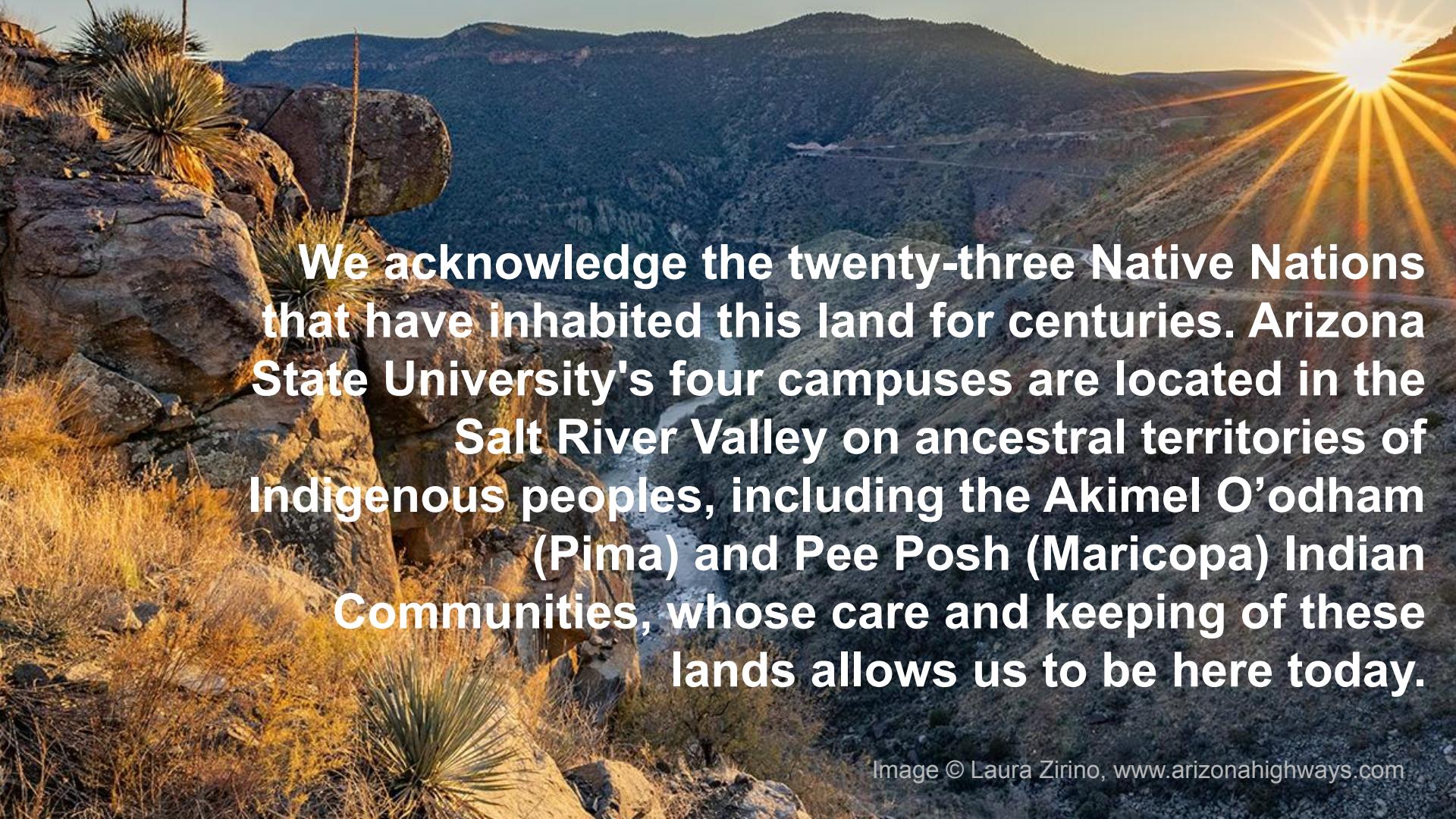
- Sam Ellis
- Ashley Ringer-Mcdonald

## Local ASU Guides & Helpers

- Edis Jakupovic
- Rekha Joshi
- Leah Repa
- Josh Uy
- Chenou Zhang



\*online only

A scenic landscape at sunset, featuring large, light-colored boulders in the foreground, some with spiky desert plants like yucca growing from them. In the background, there are rolling hills and mountains covered in green vegetation. The sun is low on the right side of the frame, casting long, golden rays across the sky and illuminating the tops of the mountains.

**We acknowledge the twenty-three Native Nations  
that have inhabited this land for centuries. Arizona  
State University's four campuses are located in the  
Salt River Valley on ancestral territories of  
Indigenous peoples, including the Akimel O'odham  
(Pima) and Pee Posh (Maricopa) Indian  
Communities, whose care and keeping of these  
lands allows us to be here today.**

# Code of Conduct



## Diversity, Equity, and Inclusion Statement

MDAnalysis strives to ensure a welcoming, inclusive space for all. As a *NumFOCUS* sponsored project, we fully support their *Diversity & Inclusion in Scientific Computing* (DISC) mission and abide by their Diversity Statement:

*"NumFOCUS welcomes and encourages participation in our community by people of all backgrounds and identities. We are committed to promoting and sustaining a culture that values mutual respect, tolerance, and learning, and we work together as a community to help each other live out these values."*

## Code of Conduct

1. Be friendly and patient
2. Be welcoming
3. Be considerate
4. Be respectful
5. Be careful in the words that you choose
6. Moderate your expectations
7. When we disagree, try to understand why
8. A simple apology can go a long way



<https://www.mdanalysis.org/pages/conduct/>

MDAnalysis Code of Conduct



Harassment or other CoC violations? Report!

# Ombudspersons



**Need to talk to someone or something doesn't feel right?**



**Oliver Beckstein** (at workshop)

- email: <[orbeckst@mdanalysis.org](mailto:orbeckst@mdanalysis.org)>
- discord: @orbeckst
- phone: **+1 480 727 9765** (no SMS)



**Irfan Alibay** (at workshop and online)

- email: <[ialibay@mdanalysis.org](mailto:ialibay@mdanalysis.org)>
- discord: @highspeedmode



**Jenna Swarthout Goddard** (online)

- email: <[community@mdanalysis.org](mailto:community@mdanalysis.org)>
- discord: @jennaswa



# Health and Safety



Lorenzo Casalino, Amaro Lab, U.C. San Diego

## COVID19 – risk differs for everyone, so we want to make it possible for everyone to attend

- Please do **not attend if you are testing positive for SARS-CoV-2**
- Limited number of test kits at registration — ask if you need one.
- Testing positive during workshop?
  - Please stay at your hotel.
  - Inform organizers or [workshops@mdanalysis.org](mailto:workshops@mdanalysis.org)
  - We will inform attendees about positive cases (*anonymously*)
- Masks are not required but welcome — we support your choice.
  - Colorful masks available at registration.



# Fire evacuation plan



## ASSEMBLY POINTS

### 1 – Primary Assembly Point

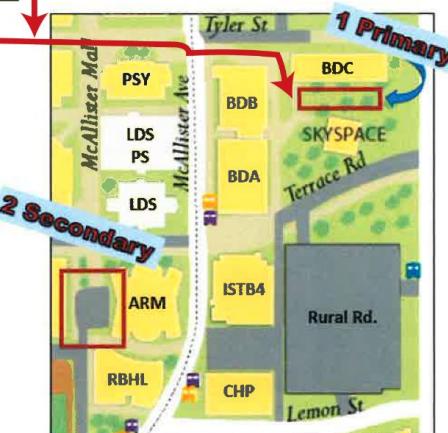
Exit the building and make your way towards the south side of BioDesign Building C (near BDC on map to the right)

### 2 – Secondary Assembly Point

Exit the building and make your way towards the parking lot area west of the ARM building (see map on right)

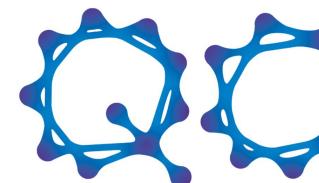
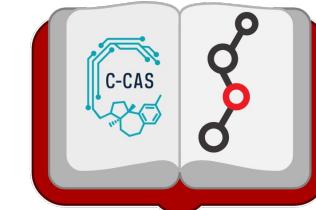
#### IMPORTANT REMINDERS

- Dial 911 to report the emergency
- Use **STAIRS**, do not use **ELEVATORS!**
- Assist those with special needs
- Go to *Primary Assembly Point* (if Primary is inaccessible, proceed to Secondary)
- Help floor wardens take roll
- Do not RE-ENTER building until instructed to do so!



# What is MolSSI?

- The **Molecular Science Software Institute** (the MolSSI) is an NSF-funded institute based at Virginia Tech (USA). It aims to be a worldwide nexus of science, education, and collaboration for the world-wide computational molecular sciences community.
  - Develop software infrastructure and provide software expertise
  - Offer education and training for scientists at all levels of their career
  - Provide community engagement and leadership.



Since 2017, over 2000 students have participated MolSSI Education workshops.

on,

## Our curriculum

Programming languages and tools...



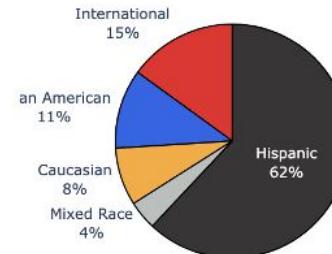
...in the context of computational molecular science.

Online Lessons

<http://education.molssi.org/resources>



MolSSI Education engages audiences who are...  
geographically and demographically diverse



Student demographics for the 2021 MolSSI Workshop in collaboration with the Tapia Center at Rice University. Students in this program were 58% female.

## Partnerships and Collaboration



Partnership for Research and Education in Chemistry  
PATHWAY TO DIVERSITY PROGRAM



RICE ENGINEERING  
Richard Tapia Center for Excellence and Equity

[education.molssi.org](http://education.molssi.org)

# What is MDAnalysis?



## Mission

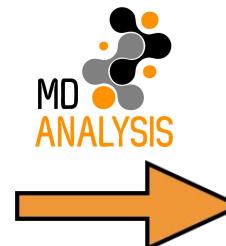
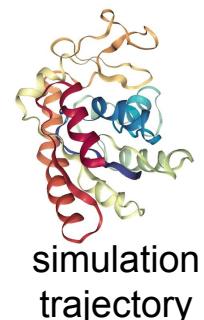
The MDAnalysis **Community** is interested in all facets of **working with data in the computational molecular sciences**. We welcome everyone. **We all follow our Code of Conduct** and strive to create an environment that is **welcoming to all**. Our primary purpose is to produce software that scientists in academia and industry will trust to use in their research.



We develop and maintain projects related to the broader goal of processing and analyzing data in the computational molecular sciences. We aim to empower users/developers to work with our packages following **FAIR principles**. Our central package is the **MDAnalysis library** for the analysis of computer simulations of many-body systems at the molecular scale.

We believe that scientific software should be open to all while using best practices to maintain high standards of correctness and reproducibility. We emphasize educating our users to make best use of the tools that we produce, to enable them to become contributors to our community and code bases.

<https://www.mdanalysis.org/>



“accessible”  
structured data  
`numpy.ndarray()`

```
import MDAnalysis as mda
u = mda.Universe(TOPOL, TRAJ)
p = u.select_atoms("protein")

cog_xy = p.positions[
    :, :2].mean(axis=0)
```



# What is MDAnalysis?



**User Group Meeting 2024**  
London, August 21-23

forums (discord, GH discussions), X, bluesky, LinkedIn

community



## workshops

OUTREACHY



STATION  
SOCIALLY-DIRECTED SCIENCE AND TECHNOLOGY

Google  
Summer of Code

education & outreach

eco-system



MDAKits

[mdakits.mdanalysis.org](http://mdakits.mdanalysis.org)

interoperability  
(RDKit, blender, ...)



MDAnalysis  
library  
(MDAnalysis)

other software (data processing in molecular sciences)

[GitHub.org/MDAnalysis/mdanalysis](https://GitHub.org/MDAnalysis/mdanalysis)

open source

[GitHub.org/MDAnalysis](https://GitHub.org/MDAnalysis)

# Thank you, MDAnalysis community!



**205 code contributors and countless community members**

Naveen Michaud-Agrawal, Elizabeth J. Denning, **Oliver Beckstein**, Danny Parton, Philip Fowler, **Tyler Reddy**, Joseph Goose, **Jan Domanski**, Benjamin Hall, Paul Rigor, David Caplan, Christian Beckstein (logo), **Sébastien Buchoux**, Joshua L. Adelman, Lukas Grossar, Andy Somogyi, Lukas Stelzl, Jinju Lu, Joshua L. Phillips, Zhuyi Xue, Xavier Deupi, **Manuel Nuno Melo**, Robert McGibbon, **Richard J. Gowers**, Alejandro Bernardin, Lennard van der Feltz, Matthieu Chavent, Joe Jordan, Alex Nesterenko, Caio S. Souza, Sean L. Seyler, **David L. Dotson**, Carlos Yanez S., Kyle J. Huston, Isaac Virshup, **Max Linke**, Gorman Stock, **Jonathan Barnoud**, Hai Nguyen, Balasubramanian, Mattia F. Palermo, Utkarsh Saxena, Abhinav Gupta, John Detlefs, Eugen Hruska, Bart Bruininks, **Fiona B. Naughton**, Robert Delgado, Wouter Boomsma, **Matteo Tiberti**, Tone Bengtsen, Shantanu Srivastava, Pedro Reis, Ruggero Cortini, Zhiyi Wu, Kashish Punjani, Utkarsh Bansal, Shobhit Agarwal, Vedant Rathore, Akshay Gupta, Juan Eiros Zamora, Jon Kapla, Sang Young Noh, Andrew William King, Kathleen Clark, Dominik 'Rathann' Mierzejewski, Nestor Wendt, **Micaela Matta**, Jose Borreguero, Sören von Bülow, Nabarun Pal, Mateusz Bieniek, Paul Smith, Navya Khare, **Johannes Zeman**, Ayush Suhane, Davide Cruz, Shujie Fan, Andrew R. McCluskey, Henry Mull, **Irfan Alibay**, **Philip Loche**, Matthew W. Thompson, Ali Ehlen, Daniele Padula, Ninad Bhat, Fenil Suchak, Yibo Zhang, Luís Pedro Borges Araújo, Abhishek A. Kognole, **Rocco Meli**, **Lily Wang**, Matthijs Tadema, Joao Miguel Correia Teixeira, Charlie Cook, Yuanyu Chang, Guillaume Fraux, Ivan Hristov, Michael Quevillon, Hao Tian, **Hugo MacDermott-Opeskin**, Anshul Angaria, Shubham Sharma, **Yuxuan Zhuang**, Cédric Bouyssat, Abhishek Shandilya, Morgan L. Nance, Faraaz Shah, Wiep van der Toorn, Siddharth Jain, Ameya Harmalkar, Shakul Pathak, Andrea Rizzi, William Glass, Marcello Segà, Edis Jakupovic, Nicholas Craven, Mieczyslaw Torchala, Ramon Crehuet, Haochuan Chen, Karthikeyan Singaravelan, Ian Aditya Kamath, Leonardo Barneschi, Henrik Jäger, Jan Stevens, Orion Cohen, Dimitrios Papageorgiou, Hannah Pollak, Estefania Barreto-Ojeda, Paarth Thadani, Henry Kobil, Kosuke Kudo, Sulay Shah, Alexander Yang, Filip T. Szczypiński, Marcelo C. R. Melo, Mark D. Driver, Kevin Boyd, Atharva Kulkarni, Yantong Cai, Bjarne Feddersen, Pratik Gupta, Alexander Gorfer, Aya M. Alaa, Kazi Shudipto Amin, Alia Lescoulie, Henok Ademtew, Uma D Kadam, Tamandeep Singh, Mingyi Xue, Meghan Osato, Anirvinya G, Rishabh Shukla, Manish Kumar, Aditi Tripathi, Sukeerti T, Kavya Bisht, Mark Verma, Marcelo D. Poletto, Ricky Sexton, Rafael R. Pappalardo, Tengyu Xie, Raymond Zhao, Haleema Khan, Jennifer A Clark, Jake Fennick, Utsav Khatu, Patricio Barletta, Mikhail Glagolev, Christian Pfaendner, Pratham Chauhan, Meet Brijwani, Vishal Parmar, Moritz Schaeffler, Xu Hong Chen, Domenico Marson, Ahmed Salah Ghoneim, Alexander Schlaich, Josh Vermaas, Xiaoxu Ruan, Egor Marin, Shaivi Malik, Daniel J. Evans, Mohit Kumar, Shubham Kumar, Zaheer Timol, Geongi Moon, Sumit Gupta, Heet Vekariya, Lawson Woods, Johannes Stöckelmaier, **Jenna M. Swarthout Goddard**, Aditya Keshari, Philipp Stärk, Kai Niklas Spauszus, Sampurna Mukherjee, Leon Wehrhan, Valerij Talagayev



# Housekeeping

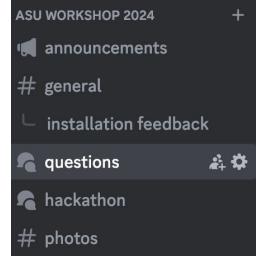
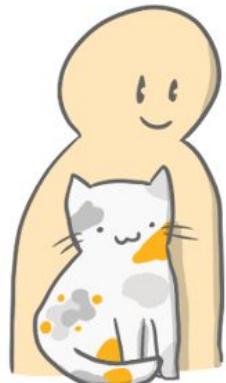


- Sessions will be **recorded** and parts may be posted <https://www.youtube.com/@mdanalysis3040>
- Online: **captions** are available online by clicking *More → Captions → Show Captions*
- Ask **questions**
  - Online: Ask questions on **discord** in the relevant *questions* post
  - In-person: Raise your hand or  
use the **sticky note system**  
or on **discord**.



<https://discord.gg/dMQWjNcZmh>

Join the MDAnalysis Discord  
server



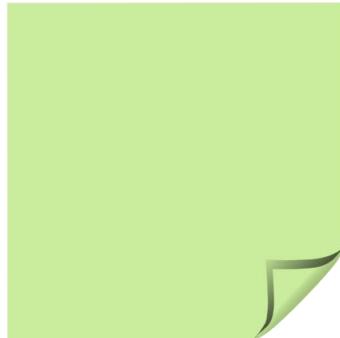
Cats by [@ExplainedByCats](#)  
(Fiona Naughton)

# Using the Sticky Notes



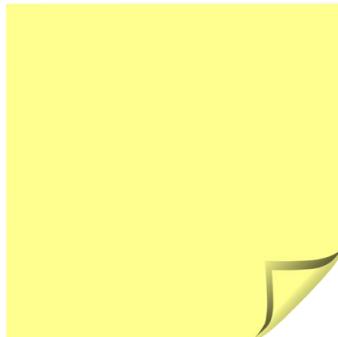
## **GREEN** sticky note:

Ready to move on /  
All good



## **YELLOW** sticky note:

I could probably use a pause /  
some thoughts



## **ORANGE** sticky note:

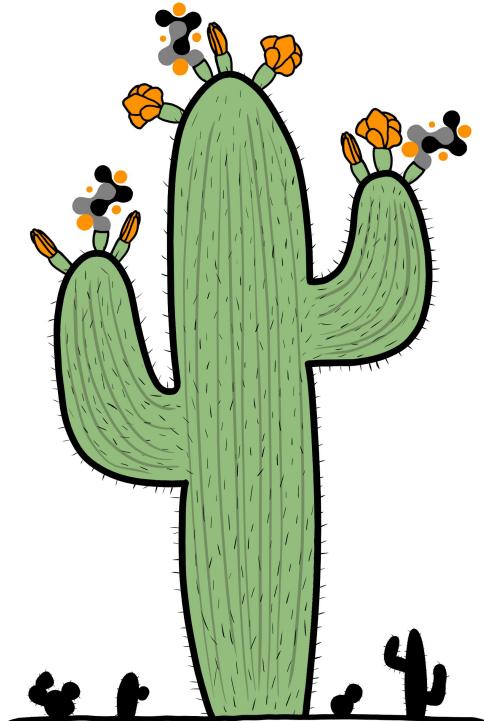
Please stop! /  
Ahh, I need help!



# Schedule - Day 1



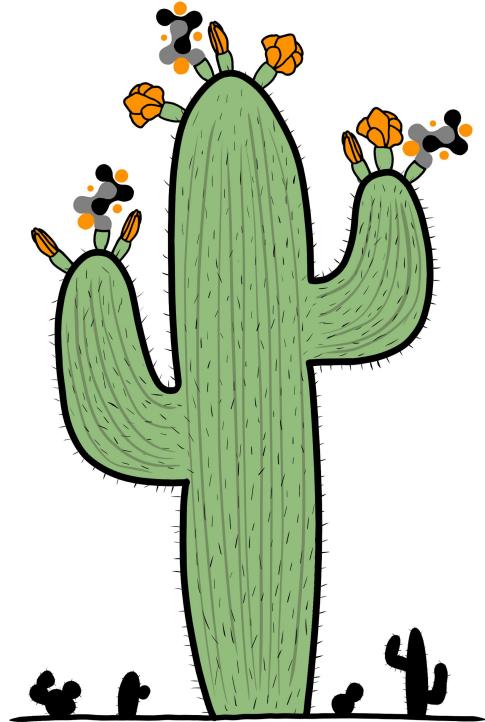
- **08:00 - 09:00 MST:** *Breakfast and registration*
- **09:00 - 09:15 MST:** Introduction
- **09:15 - 11:00 MST:** Introduction to MDAnalysis
  - 09:15 - 10:05 MST: Universe and Atoms
  - 10:05 - 11:00 MST: File I/O and working with trajectories
- **11:00 - 11:15 MST:** Break
- **11:15 - 12:45 MST:** MDAnalysis
  - 11:15 - 12:05 MST: Analysis
  - 12:05 - 12:45 MST: Advanced Tips and Tricks
- **12:45 - 12:55 MST:** Introduction to MDAKits
- **12:55 - 14:00 MST:** *Lunch + Group Photo*
- **14:00 - 14:30 MST:** Introduction to the Bash Shell
- **14:30 - 16:00 MST:** Python Packaging
- **16:00 - 16:15 MST:** Break
- **16:15 - 18:15 MST:** Version Control with git and GitHub



# Schedule - Day 2



- **08:00 - 09:00 MST:** *Breakfast*
- **09:00 - 10:45 MST:** Testing and pytest
- **10:45 - 11:00 MST:** *Break*
- **11:00 - 12:30 MST:** Documentation
- **12:30 - 13:30 MST:** *Lunch*
- **13:30 - 13:40 MST:** Hackathon Introduction
- **13:40 - 15:30 MST:** Hackathon
- **15:30 - 15:45 MST:** *Break*
- **15:45 - 17:15 MST:** Hackathon
- **17:15 - 17:45 MST:** Hackathon Wrap-Up and Showcase
- **17:45 - 18:00 MST:** Closing Remarks



# Setting up Your Environment



- **Materials** on ‘jun24-ws’ branch in  
<https://github.com/MDAnalysis/MDAnalysisMolSSIWorkshop-Intermediate2Day> repo
- Full **installation** instructions can be found in [INSTALL.md](#)
- Make sure to **pull the repo** again to ensure all materials are up-to-date!

```
$ cd ~  
$ git clone  
https://github.com/MDAnalysis/MDAnalysisMolSSIWorkshop-Intermediate2Day.git  
  
$ cd ~/MDAnalysisMolSSIWorkshop-Intermediate2Day  
$ git pull
```



GitHub repository

<https://github.com/MDAnalysis/MDAnalysisMolSSIWorkshop-Intermediate2Day>

And now...



**...on to the first session!**



# Final Announcements



- Sessions will be **recorded** and parts may be posted on  
<https://www.youtube.com/@mdanalysis3040>
- We will send a follow-up email with all relevant **resources**
- We want to hear **your feedback!** Please complete the survey.
- More workshops to come! For updates follow our:
- **Discord:** <https://discord.gg/dMOWjNcZmh>
- **GitHub Discussions:** <https://github.com/MDAnalysis/mdanalysis/discussions>
- **Blog:** <https://www.mdanalysis.org/blog/>
- **LinkedIn:** <https://linkedin.com/company/mdanalysis>



<https://forms.gle/ZhqBEzXfw6SqGaQ96>

Post-workshop survey

Discord invite

