



MDAnalysis Toolkits

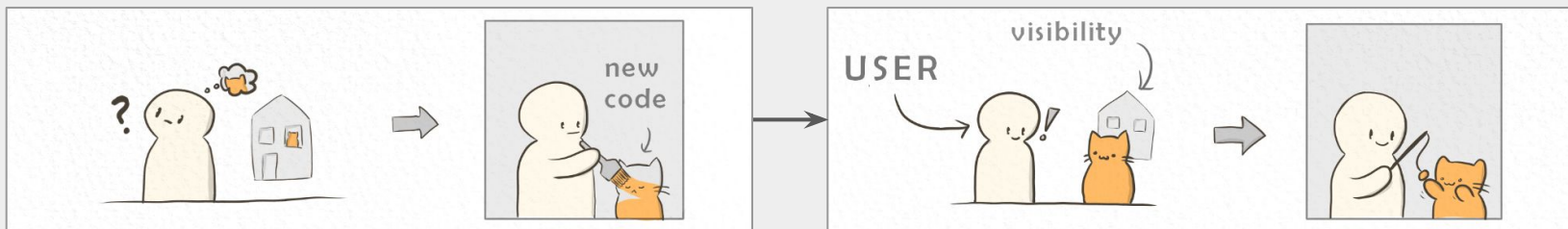
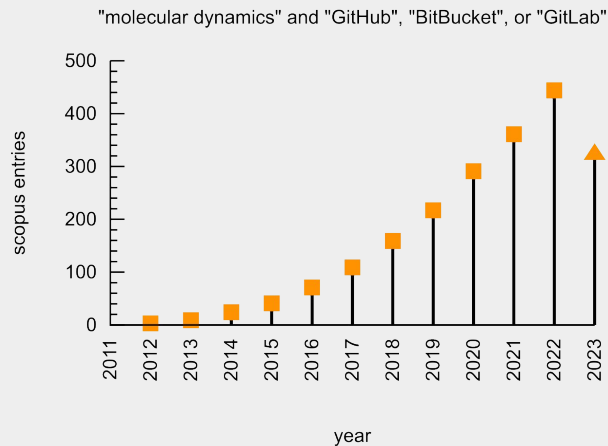
A framework for FAIR-compliant molecular simulation analysis

Irfan Alibay, Lily Wang, Fiona Naughton, **Ian Kenney**, Jonathan Barnoud, Richard J Gowers, Oliver Beckstein

Sharing scientific code



- Reduced barrier for hosting and sharing code
- Share methods and analyses
 - Reduced duplication effort



Cat illustrations by @explainedbycats (Fiona Naughton)

FAIR principles for research software



Findable

Accessible

Interoperable

Usable and Reusable

scientific data

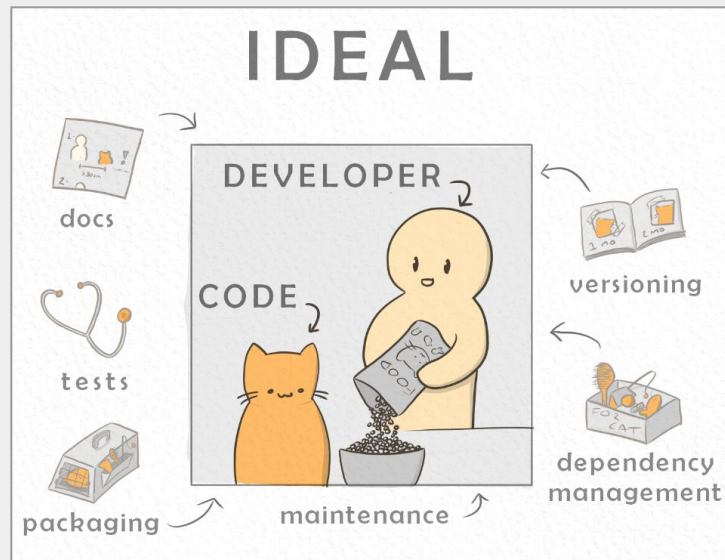
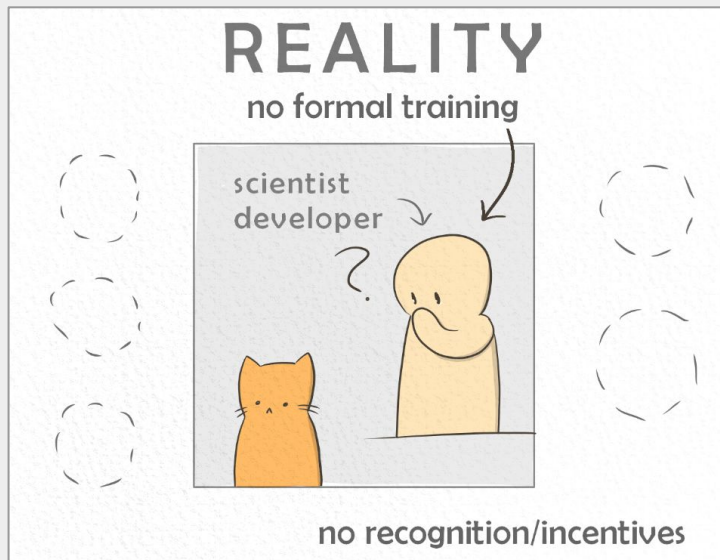
OPEN ARTICLE

Introducing the FAIR Principles for research software

Check for updates

Michelle Barker^{1,✉}, Neil P. Chue Hong², Daniel S. Katz³, Anna-Lena Lamprecht⁴, Carlos Martinez-Ortiz⁵, Fotis Psomopoulos⁶, Jennifer Harrow⁷, Leyla Jael Castro⁸, Morane Gruenpeter⁹, Paula Andrea Martinez¹⁰ & Tom Honeyman¹¹

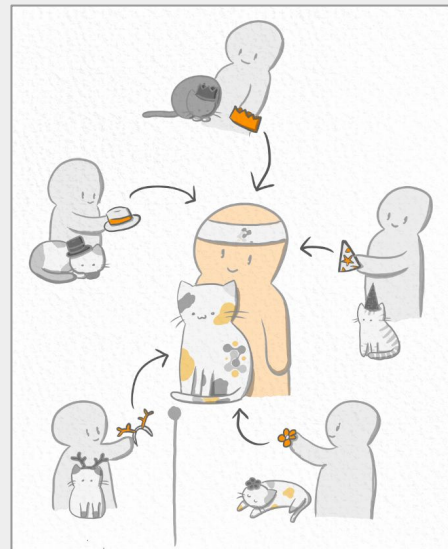
Scientists are rarely software engineers



Contributing to a mature project



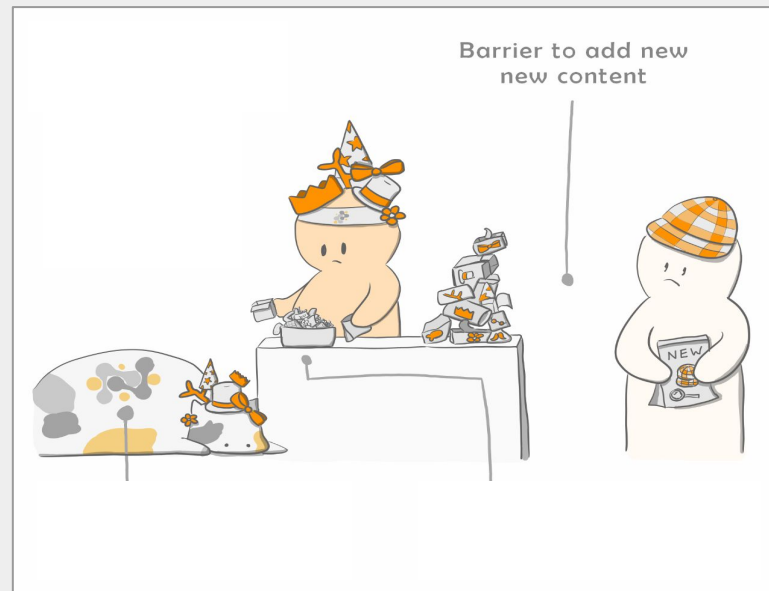
- Immediate promotion to the community
- Documentation
- Testing
- Access to experienced developers
- Potentially reduced individual effort
 - Coordinating new feature contributions
 - Fixing bugs



When and why this can fail



- Slow adoption process
- Limited method expertise amongst package maintainers
- Unmaintained modules for niche methods
- Ever-increasing code footprint for the core library



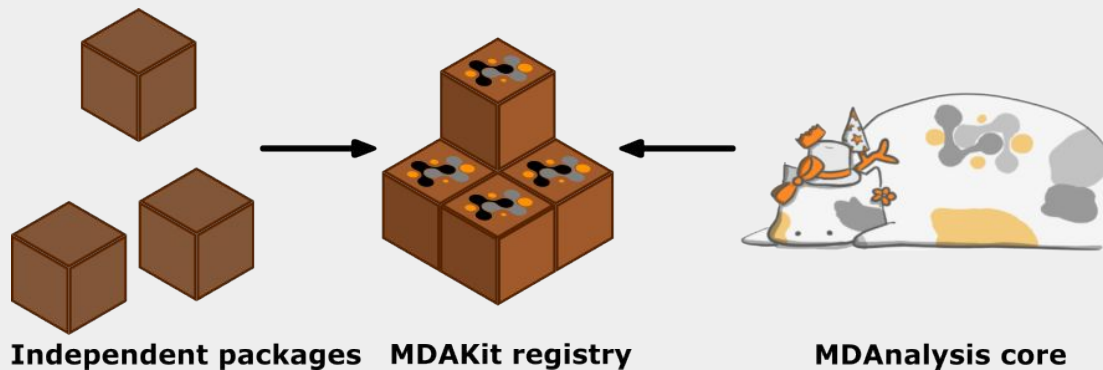
An MDAnalysis toolkit (MDAKit) framework



- **Any code that uses MDAnalysis**
- **Open source with an OSI approved license**
- Code is versioned and accessible in a version controlled repository
- Authors and maintainers are designated
- Basic documentation is provided
- Basic tests are provided



An MDAnalysis toolkit (MDAKit) framework



registry



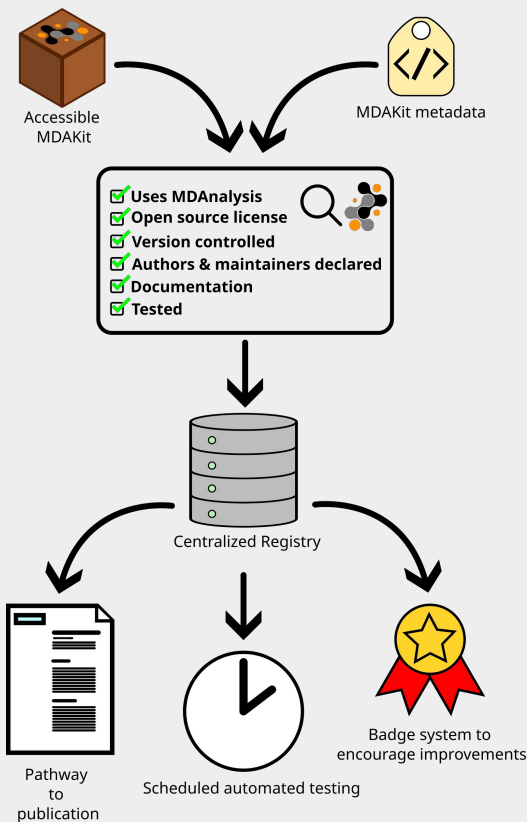
<https://mdakits.mdanalysis.org/mdakits.html>

MDAKit registration



- Visibility
- Biweekly testing of all kits
- Automated maintainer alerts on failures
- Streamlined publication with JOSS*

```
project_name: hole2-mdakit
authors:
  - https://github.com/MDAnalysis/hole2-mdakit/blob/main/AUTHORS.md
maintainers:
  - ianmkenney
description:
  A Python interface for the HOLE suite tools to analyze an ion channel
  ↳ pore or transporter pathway as a function of time or arbitrary order
  ↳ parameters.
keywords:
  - pores
  - ion channels
  - transporters
  - HOLE
license: GPL-3.0
project_home: https://github.com/MDAnalysis/hole2-mdakit
documentation_home: https://www.mdanalysis.org/hole2-mdakit/
documentation_type: API
```



Making an MDAKit



- Using the cookiecutter template
 - Automatic generation of metadata, community infrastructure, and documentation
 - Package boilerplate
 - Testing framework
 - Accompanying tutorial
- From scratch

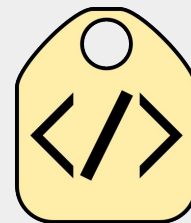
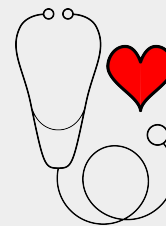
cookiecutter



tutorial



<https://github.com/MDAnalysis/cookiecutter-mdakit>
<https://mdakits.mdanalysis.org/makingakit.html>



Current selection of MDAKits



- Find the current list of MDAKits on <https://mdakits.mdanalysis.org>
- Ranging from analysis methods to user interfaces to MDAnalysis itself!

registry



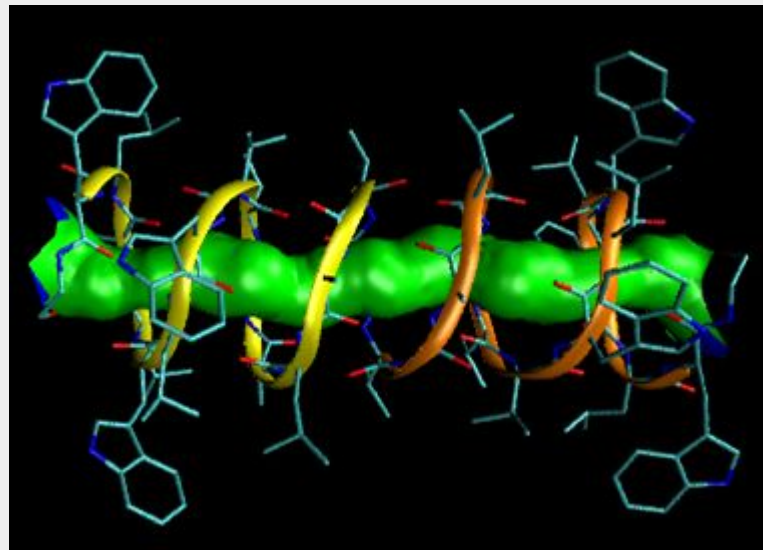
<https://mdakits.mdanalysis.org/mdakits.html>

MDAKit	Keywords	Authors	CI badges
hole2-mdakit	pores, ion channels, transporters, HOLE	hole2-mdakit authors	latest develop passed passed
lipyds	lipids, membranes	tilyminium	latest develop passed unavailable
mda-tui	transformations, TUI	mda-tui authors	latest develop passed passed
mdachecker	structure analysis, conformation checks	IAlibay	latest develop passed passed
mdaccli	command line interface, molecular-dynamics	mdaccli authors	latest develop passed passed
mdaencore	ensemble, similarity, covariance, PCA	mdaencore authors	latest develop passed passed
openmm-mdanalysis-reporter	OpenMM, Reporters	openmm-mdanalysis-reporter authors	latest develop passed passed
ProLIF	drug-design, cheminformatics, molecular-dynamics	Cédric Bouysset	latest develop passed passed
propkatraj	pKa, protein	propkatraj authors	latest develop passed passed
solvation-analysis	chemistry, electrolytes, solvation structure	orionarcher, hmadope, laurlee, IAlibay	latest develop passed passed

MDAKit showcase: mdahole2



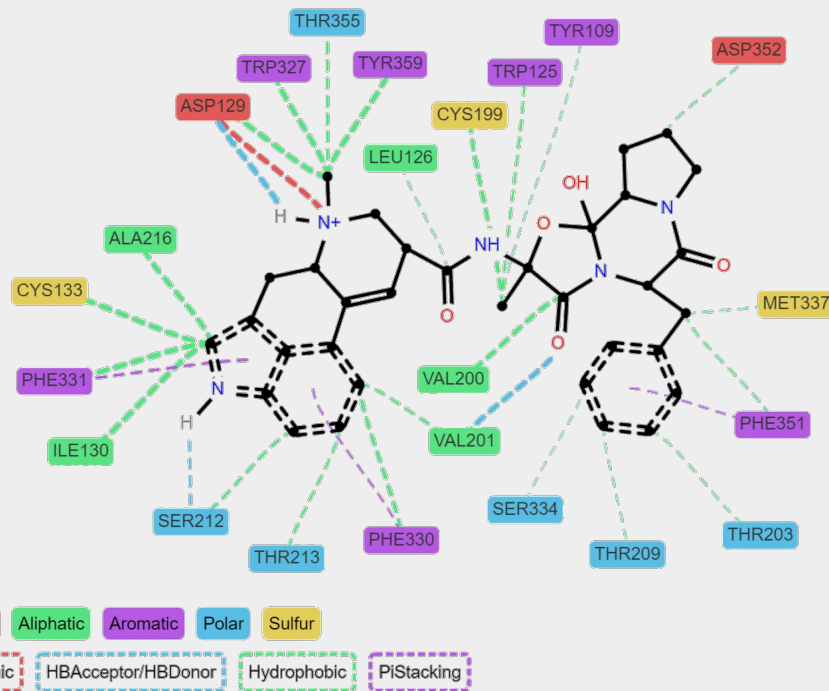
- Analysis of pathways through molecular channels
- Migrated core library specialized analysis
- Depends on external programs
 - Binaries made available on conda-forge



MDAKit showcase: ProLIF



- Generate interaction fingerprints for molecular complexes from simulation trajectories or experimental structures
- Package from Cédric Bouysset
- Code predating MDAKits



Journal of
Cheminformatics



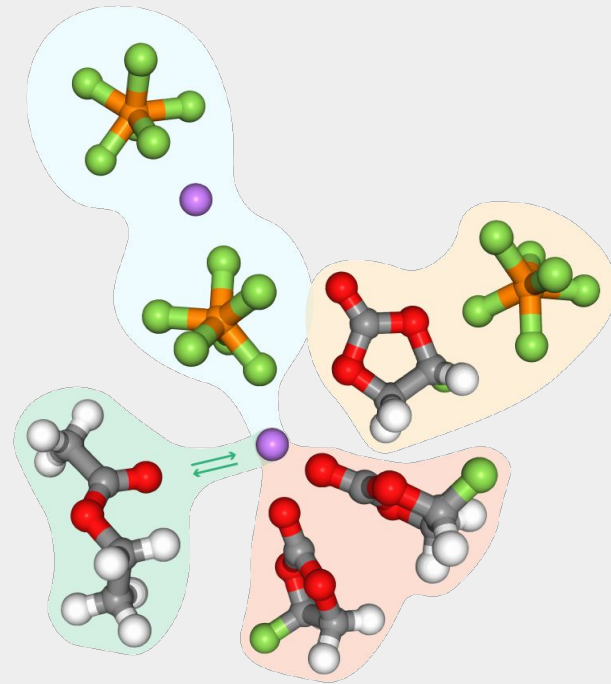
Bouysset, C., & Fiorucci, S. (2021). ProLIF: A library to encode molecular interactions as fingerprints. Journal of Cheminformatics, 13(1), 72. <https://doi.org/10.1186/s13321-021-00548-6>

Modified image from publication.

MDAKit showcase: SolvationAnalysis



- Implementation of routinely repeated analyses for solvated systems
- GSoC project from Orion Cohen
- JOSS publication



JOSS paper



Cohen, O. A., Macdermott-Opeskin, H., Lee, L., Hou, T., Fong, K. D., Kingsbury, R., Wang, J., & Persson, K. A. (2023). SolvationAnalysis: A Python toolkit for understanding liquid solvation structure in classical molecular dynamics simulations. Journal of Open Source Software, 8(84), 5183. <https://doi.org/10.21105/joss.05183>

Modified image from publication.

Conclusions



- Promoting an environment where sharing with the community is substantially easier
- Emphasis on long-term maintenance of code
- Read the SciPy Proceedings paper for more information!

SciPy paper



Alibay, I., Wang, L., Naughton, F., Kenney, I., Barnoud, J., Gowers, R., & Beckstein, O. (2023). MDAKits: A Framework for FAIR-Compliant Molecular Simulation Analysis. 76–84. <https://doi.org/10.25080/gerudo-f2bc6f59-00a>

Consider attending the MDAKits Hackathon



- Get involved by contributing an MDAKit at the Hackathon!
- Work on a suggested Kits or bring your own idea
 - Check out the project ideas in the MDAKits Hackathon section
- Hit the ground running and make sure to read the MDAKits tutorial beforehand!

MDAKits
tutorial



<https://mdakits.mdanalysis.org/makingakit.html>

Acknowledgements



188 code contributors and countless community members

Naveen Michaud-Agrawal, Elizabeth J. Denning, 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, 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, Hai Nguyen, Balasubramanian, Mattia F. Palermo, Utkarsh Saxena, Abhinav Gupta, John Detlefs, Eugen Hruska, Bart Bruininks, 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, **Philip Loche**, Matthew W. Thompson, Ali Ehlen, Daniele Padula, Ninad Bhat, Fenil Suchak, Yibo Zhang, Luís Pedro Borges Araújo, Abhishek A. Kognole, **Rocco Meli**, 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 Bouysset, Abhishek Shandilya, Morgan L. Nance, Faraaz Shah, Wiep van der Toorn, Siddharth Jain, Ameya Harmalkar, Shakul Pathak, Andrea Rizzi, William Glass, Marcello Segal, Edis Jakupovic, Nicholas Craven, Mieczyslaw Torchala, Ramon Crehuet, Haochuan Chen, Karthikeyan Singaravelan, Ian M. Kenney, Aditya Kamath, Leonardo Barneschi, Henrik Jäger, Jan Stevens, Orion Cohen, Dimitrios Papageorgiou, Hannah Pollak, Estefania Barreto-Ojeda, Paarth Thadani, Henry Kobin, 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 Ademteu, 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 Khatau, 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





Thank you! Questions?

Links



- Registry – <https://mdakits.mdanalysis.org/mdakits.html>
- Cookiecutter – <https://github.com/MDAnalysis/cookiecutter-mdakit>
- Written tutorial – <https://mdakits.mdanalysis.org/makingakit.html>
- Video tutorial – <https://www.youtube.com/watch?v=viCPUHkgSxg>