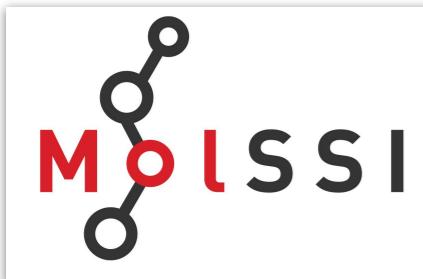


The Molecular Sciences Software Institute

... a nexus for science, education, and cooperation for the global computational molecular sciences community.

What is the MolSSI?

- Launched August 1st, 2016, funded by the National Science Foundation.
- Collaborative effort by Virginia Tech (TDC), Rice U. (C. Clementi), Stony Brook U. (R. Harrison), U.C. Berkeley (T. Head-Gordon), Stanford U. (V. Pande), Rutgers U. (S. Jha), U. Southern California (A. Krylov), and Iowa State U (T. Windus).
- Part of the NSF's commitment to the White House's National Strategic Computing Initiative (NSCI).
- Total budget of \$19.42M for five years, potentially renewable to ten years.
- Joint support from numerous NSF divisions: Advanced Cyberinfrastructure (ACI), Chemistry (CHE), and Division of Materials Research (DMR)
- Designed to **serve** and **enhance** the software development efforts of the broad field of computational molecular science.



Computational Molecular Sciences (CMS)

- The history of CMS – the sub-fields of quantum chemistry, computational materials science, and biomolecular simulation – reaches back decades to the genesis of computational science.
- CMS is now a “full partner with experiment”.
- For an impressive array of chemical, biochemical, and materials challenges, our community has developed simulations and models that directly impact:
 - Development of new chiral drugs;
 - Elucidation of the functionalities of biological macromolecules;
 - Development of more advanced materials for solar-energy storage, technology for CO₂ sequestration, etc.

CMS Codes Are Developed and Used Globally



CMS Codes Are Developed and Used Globally



These codes represent **decades** of development by thousands of programmers, and are used by **hundreds of thousands** of scientists worldwide.



Code Complexity and Historical Legacy

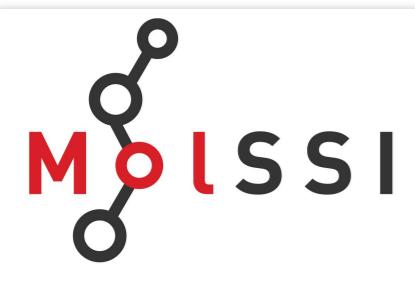
- CMS programs contain **millions of lines of hand-written code** and require **hundreds of programmers** to develop and maintain.
- Incredible **language diversity**: F77, F90, F95, HPF, C, C++, C++11, C++14,C++17, Python, perl, Javascript, etc.
- Incredible **algorithmic diversity**: structured and unstructured grids, dense and sparse linear algebra, graph traversal, fast Fourier transforms, MapReduce, and more.
- The packages have evolved in an ad hoc manner **over decades** because of the intricacy of the scientific problems they are designed to solve.

Rapidly Evolving Computing Hardware

- **Multi- and many-core architectures** are the norm, but many CMS codes are developed with limited view to parallel task management.
- Reduced-power solutions will also require improved **error recovery and checkpointing** at the software level – capabilities absent in nearly all CMS codes.
- Anticipated architectural innovations will yield **even greater hardware complexity** – more advanced accelerators, specialized computing cores, reconfigurable logic...
- Many CMS codes (especially for quantum chemistry) are limited to shared-memory paradigms and cannot yet take advantage of GPUs or **large-scale distributed-memory systems**.

Inertia in the Scientific Education Culture

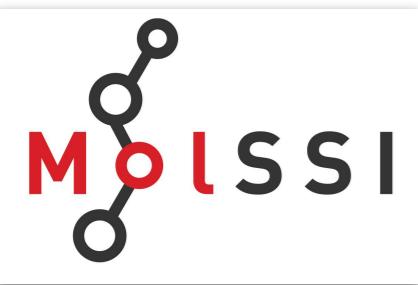
- Undergraduate programs in chemistry and physics typically require **no training in software development** or programming.
- Graduate programs in these areas require **minimal coursework** between the bachelor and Ph.D.
- Most computer science students lack the underlying knowledge of the scientific domains to help develop creative software solutions.
- **Due credit for software development** is elusive due to a culture that judges productivity based on citations of peer-reviewed papers.
- Thus, a “**just get the physics working**” approach pervades much of CMS software development.



MolSSI Goal #1

To Provide Software Expertise and Infrastructure...

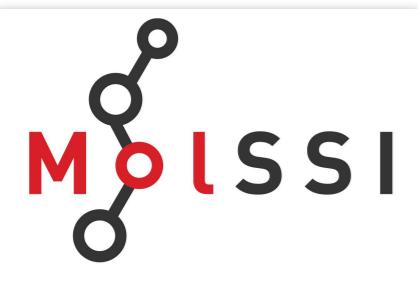
- MolSSI works with CMS research groups nationwide and internationally to design, develop, test, deploy, and maintain key code infrastructure and frameworks for the entire community.
- MolSSI interacts with partners in industry, NSF supercomputing centers, national laboratories, and international facilities to identify and act on emerging hardware trends, access leading-edge computing architectures, further educational goals, set software priorities, and identify future workforce career paths.



MolSSI Goal #2

To Provide Education and Training...

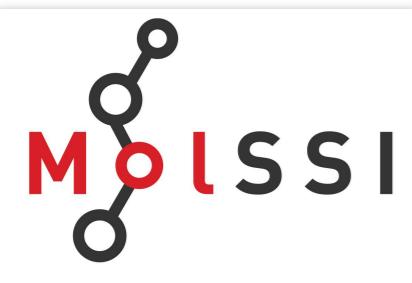
- MolSSI serves as an education and outreach nexus for the worldwide CMS community.
- MolSSI organizes summer schools, targeted workshops, high-school and undergraduate training programs, and on-line resources and classes to provide current and future CMS students with a modern and complete set of programming skills.
- MolSSI reaches beyond the traditional student cohort to computer scientists and mathematicians seeking interdisciplinary applications.
- MolSSI will deploy a Professional Master's program in Molecular Simulation and Software Engineering.



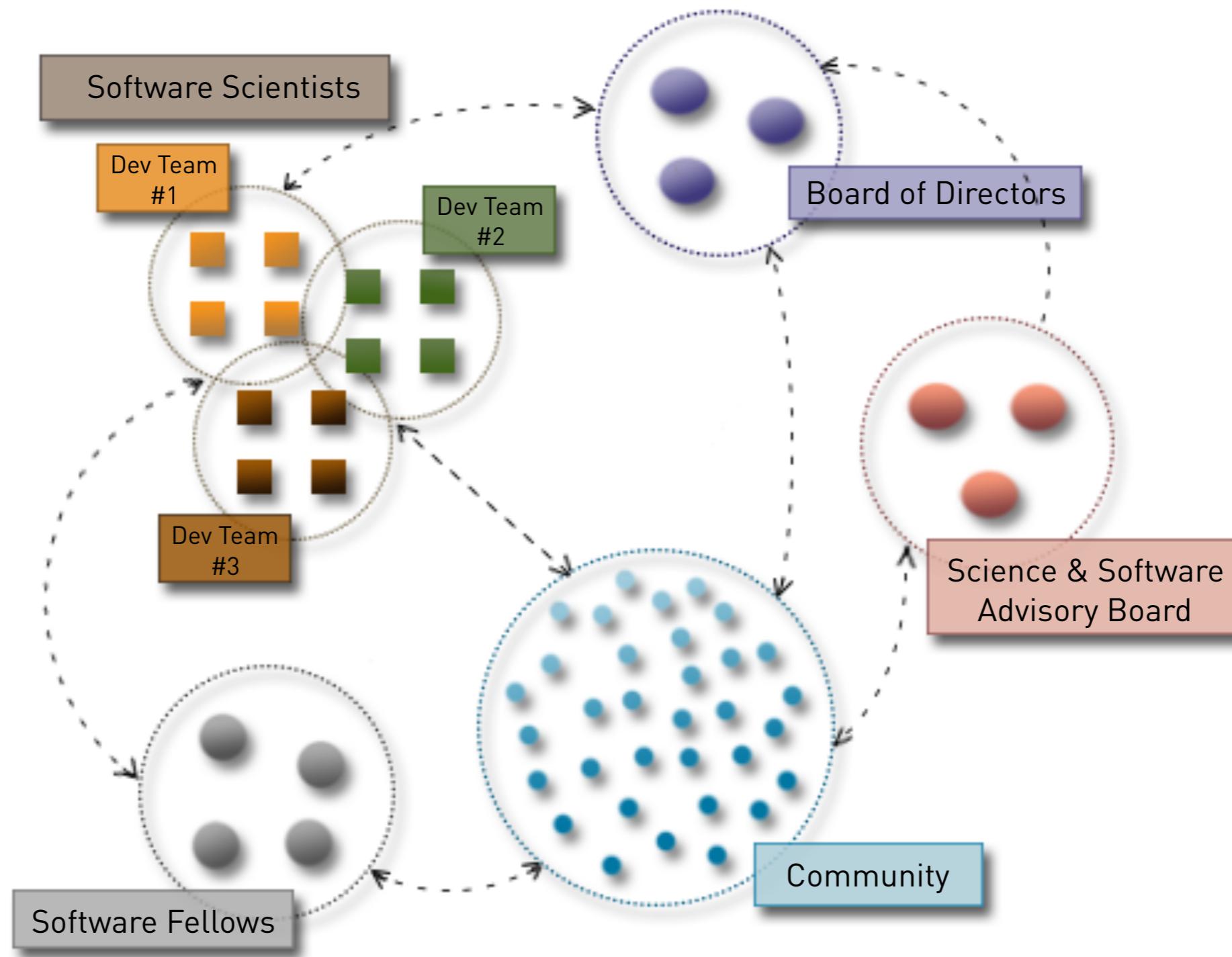
MolSSI Goal #3

To Provide Community Engagement and Leadership...

- MolSSI will enable the CMS community to establish its own standards for interoperability, best practices, and curation tools.
- Through a “**grass roots**” approach, MolSSI engages the community broadly using interoperability workshops and focus groups – and ultimately will enable the formation of a **Molecular Sciences Consortium** – to catalyze the consensus needed for standardization of data structures, APIs, and frameworks for the entire CMS software ecosystem.



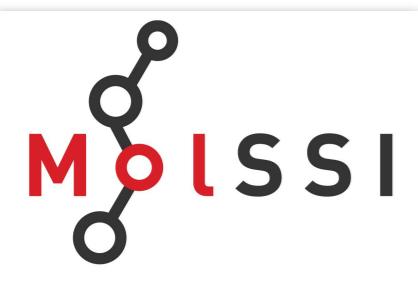
The Molecular Sciences Software Institute



MolSSI Software Scientists (MSSs)

- A team of ~12 software engineering experts, drawn both from newly minted Ph.D.s and established researchers in molecular sciences, computer science, and applied mathematics.
- Dedicated to multiple responsibilities:
 - Developing software infrastructure and frameworks;
 - Interacting with CMS research groups and community code developers;
 - Providing forums for standards development and resource curation;
 - Serving as mentors to MolSSI Software Fellows;
 - Working with industrial, national laboratory, and international partners;

Approximately 50% of the Institute's budget directly supports the MolSSI Software Scientists.



MolSSI Software Scientists



Paul Saxe
Lead Software Scientist
Ph.D., U.C. Berkeley
Materials, Molecular Dynamics,
and Quantum Chemistry

MolSSI Software Scientists



Doaa Altarawy
Ph.D., Virginia Tech
Computer Science
Machine Learning,
Software Engineering



Taylor Barnes
Ph.D., Caltech
Quantum Chemistry,
High Performance Computing,
Code Optimization

MolSSI Software Scientists



Eliseo Marin-Rimoldi
Ph.D., Univ. Notre Dame
Monte Carlo Methods, Phase
Equilibria, Statistical
Thermodynamics

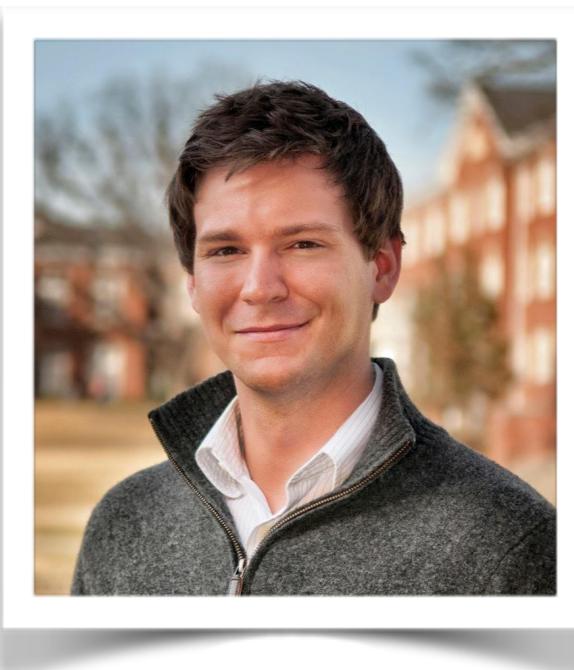


Jessica Nash
Ph.D., N.C. State Univ.
Soft Materials,
Molecular Dynamics

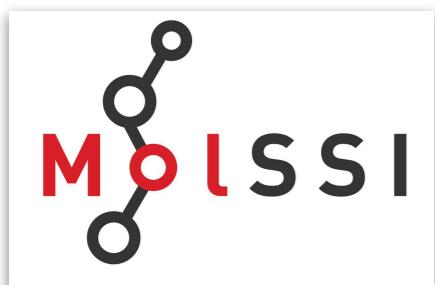
MolSSI Software Scientists



Benjamin Pritchard
Ph.D., Univ. Buffalo
Quantum Chemistry,
Interoperability Frameworks,
Microarchitecture Optimization



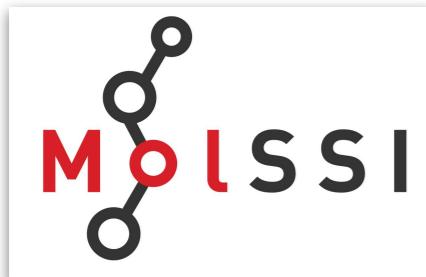
Daniel Smith
Ph.D., Auburn University
Quantum Chemistry,
Workflows,
Rapid Prototyping



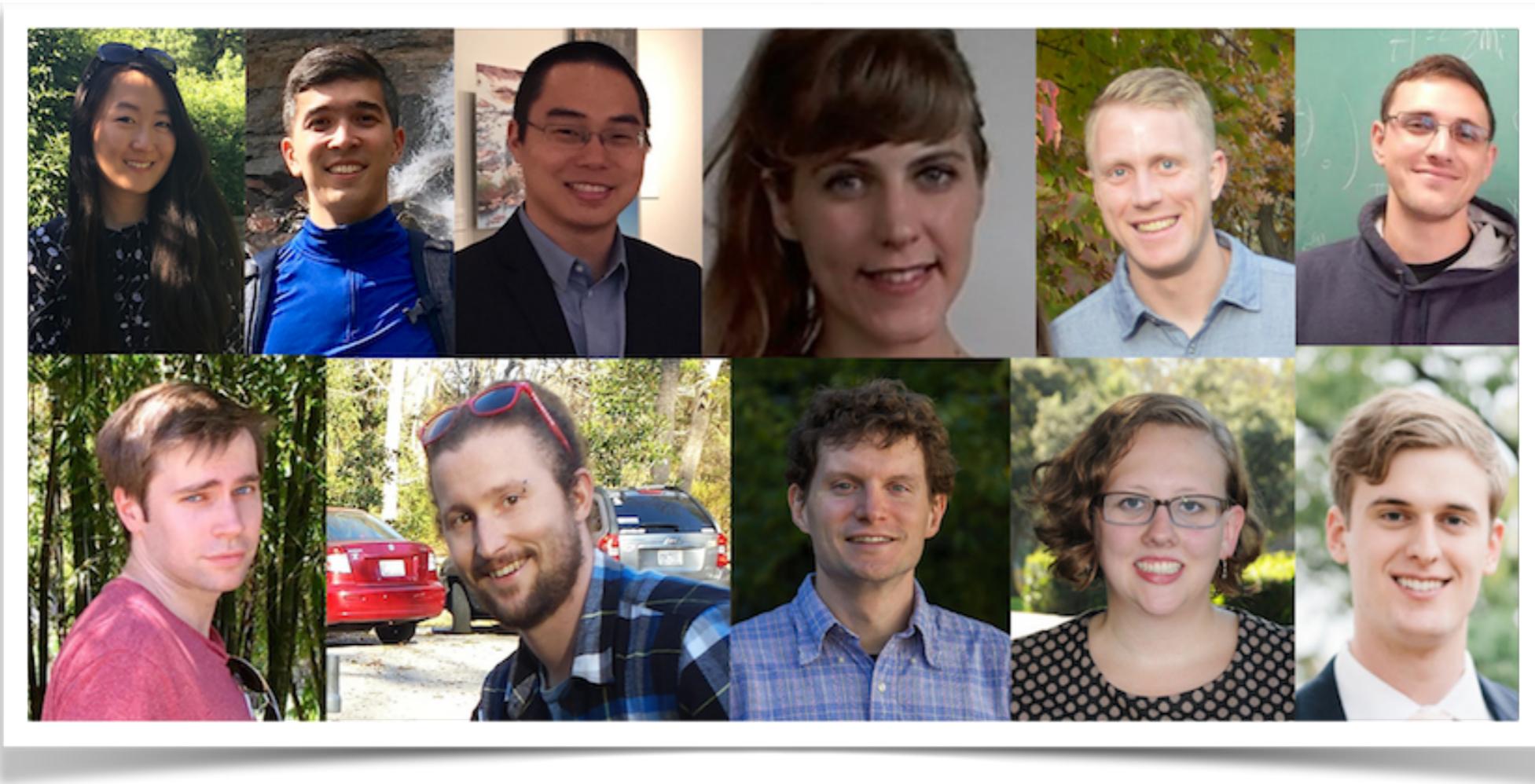
MolSSI Software Fellows (MSFs)

- A cohort of ~20 Fellows supported simultaneously – graduate students and postdocs selected by the Science and Software Advisory Board from research groups across the U.S.
- Fellows work directly with both the Software Scientists and the MolSSI Directors, thus providing a conduit between the Institute and the CMS community itself.
- Fellows work on their own projects, as well as contribute to the MolSSI development efforts where appropriate, and they engage in outreach and education activities under the Institute guidance.
- Funding for MolSSI Software Fellows follows a flexible, two-phase structure, providing up to two years of support.

Approximately 25% of the Institute's budget directly supports the MolSSI Software Fellows.



2018 Phase-I MolSSI Software Fellows



Joshua Rackers (Washington U in St. Louis)

Jocelyn Sunseri (U Pittsburgh)

Dr. Tyler Takeshita (UC Berkeley)

Ruslan Tazhigulov (Boston U)

Jing Yang (U Pennsylvania)

Adam Abbott (U Georgia)

Caitlin Banna (UC Irvine)

Dr. M. Eric Irrgang (U Virginia)

Aaron Mahler (Duke U)

Dr. James McClain (Cal Tech)

Justin Provazza (Boston U)

2018 Phase-I MolSSI Software Fellows



Joshua Rackers (Washington U in St. Louis)
Jocelyn Sunseri (U Pittsburgh)
Dr. Tyler Takeshita (UC Berkeley)
Ruslan Tazhigulov (Boston U)
Jing Yang (U Pennsylvania)

Adam Abbott (U Georgia)
Caitlin Banna (UC Irvine)
Dr. M. Eric Irrgang (U Virginia)
Aaron Mahler (Duke U)
Dr. James McClain (Cal Tech)
Justin Provazza (Boston U)

2018 Phase-II MolSSI Software Fellows



Srinivas Mushnoori
Rutgers University

Adviser: Prof. Meenakshi Dutt

“Development and extension of a flexible,
scalable Replica Exchange framework”

Kate Lebold

Penn State University

Adviser: Prof. William Noid

“Thermodynamic state point transferability of
coarse-grained models through the force-,
pressure-, and energy-matching methodologies”



2018 Phase-II MolSSI Software Fellows



Erik Thiede

University of Chicago

Advisers: Profs. Aaron Dinner and Jonathan Weare
“Protein dynamics in the absence of collective variables through dynamics-dependent statistics and the study of partial differential equations”

Marc Riera-Riambau

University of California, San Diego

Adviser: Prof. Francesco Paesani

“Development and application of many-body potential energy functions in a user friendly software interface”



2018 Phase-II MolSSI Software Fellows



David Williams-Young
University of Washington
Adviser: Prof. Xiaosong Li

“High-performance and scalable relativistic electronic structure methods for the ab initio prediction of molecular properties”

Oanh Vu

Vanderbilt University

Adviser: Prof. Jens Meiler

“Machine-learning models used for drug discovery through local similarity-based descriptors with traditional autocorrelation descriptors”



2018 Phase-II MolSSI Software Fellows

Boyi Zhang

University of Georgia

Adviser: Prof. Fritz Schaefer

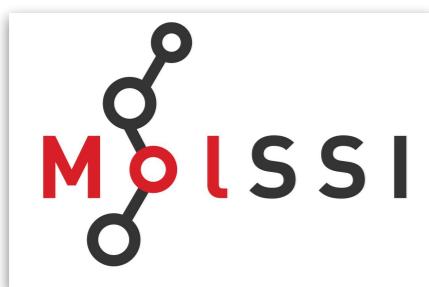
“A unified software platform that effectively implements various adaptive Quantum Mechanics/Molecular Mechanics (QM/MM) methods”



MolSSI Community Code Partners So Far...

- | | | | |
|------------|-------------|------------|--------------------|
| ● ACESIII | ● Dirac | ● Molpro | ● PARSEC |
| ● ADF | ● DL_POLY | ● MPQC | ● PCMSolver |
| ● Amber | ● ELSI | ● MRChem | ● PLUMED |
| ● APBS | ● FHI-aims | ● NAMD | ● PSI4 |
| ● BOSS | ● GAMESS | ● NWChem | ● Q-Chem |
| ● CFOUR | ● Gaussian | ● NWChemEx | ● QBox |
| ● CHARMM | ● Gromacs | ● ONETEP | ● QMworks |
| ● Columbus | ● LAMMPS | ● OpenMM | ● Quantum ESPRESSO |
| ● Dalton | ● Molcas | ● Orca | ● Schrödinger |
| ● Tiger-Cl | ● Turbomole | ● VASP | |

We encourage all community codes in the
computational molecular sciences to work with us!



MolSSI Board of Directors



Cecilia Clementi, Rice University, Co-Director for International Engagement



T. Daniel Crawford, Virginia Tech, Director



Robert J. Harrison, Stony Brook University, Co-Director for Parallel Computing and Emerging Technologies



Teresa Head-Gordon, U.C. Berkeley, Co-Director for Laboratory, Industrial, and Academic Outreach and Education



Shantenu Jha, Rutgers University, Co-Director for Software Engineering Process, Middleware, and Infrastructure



Anna Krylov, U. Southern California, Co-Director for Quantum Chemistry and Materials



Vijay Pande, Stanford University, Co-Director for Molecular Simulation



Theresa Windus, Iowa State University, Co-Director for Code and Data Interoperability

Science and Software Advisory Board



Kathy Yelick
U.C. Berkeley and LBNL
Chair

Rigoberto Hernandez
Johns Hopkins University
Vice-Chair



Alán Aspuru-Guzik
Harvard University

Nathan Baker
Pacific Northwest
National Labs



Nicholas Berente
University of Georgia

Charles Brooks
University of Michigan

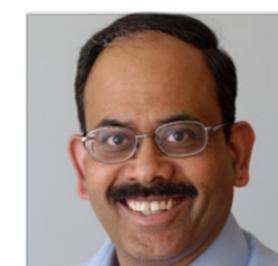


Science and Software Advisory Board



Ian Foster
University of Chicago

Laxmikant Kale
University of Illinois



Todd Martinez
Stanford University

Julia Rice
IBM Almaden



Angel Rubio
MPI – Hamburg

George Shields
Furman University



MolSSI Education Initiative

- Primary objectives:
 - Teamwork and collaborative projects
 - MolSSI Best Practices
 - Open-source paradigms
 - Modern tools for modern software
 - Teach content to increase both a students scientific capability and industrial marketability
 - The “true” cost of ownership
- Core Components:
 - Software Carpentry style courses
 - Learning groups and hands on practice
 - Taught by experts in the field

Inspiration:



MolSSI Education Initiative

- Workshop Types (2017-2019)
 - MolSSI Software Summer School (biennial @ 60 students)
 - MSF Bootcamp (2/year @ 10 students)
 - SSE/SSI Infrastructure Workshop (1/year @ 20 students)
 - CMS Novice Workshops (3-6/year @ 30-60 students)
- Upcoming events:
 - CMS Novice Workshop @ Duke (June 2018)
 - SSE/SSI Infrastrucure Workshop @ NYC (Fall 2018)
 - MSF Bootcamp @ MolSSI HQ (July 2018)
 - Summer School and Workshop - Parallel Computing in Molecular Sciences @ Berkeley (August 2018)



MolSSI Workshops: Leadership-Driven



- The Cyberinfrastructure Requirements and Challenges of Molecular Science, Rice University, October 8 & 9, 2016 (**Organizers:** Cecilia Clementi and Shantenu Jha)



- Computational Materials Software, U.C. Berkeley, February 2-4, 2017 (**Organizers:** Teresa Head-Gordon, Steven Louie, Jeff Neaton)



- Biomolecular Simulation, Stanford University, April 6 & 7, 2017 (**Organizer:** Vijay Pande)

MolSSI Workshops: Leadership-Driven

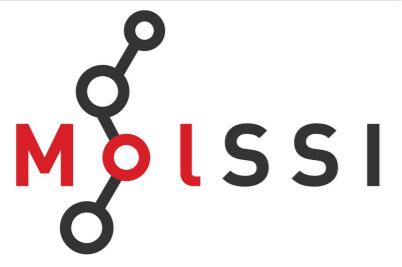
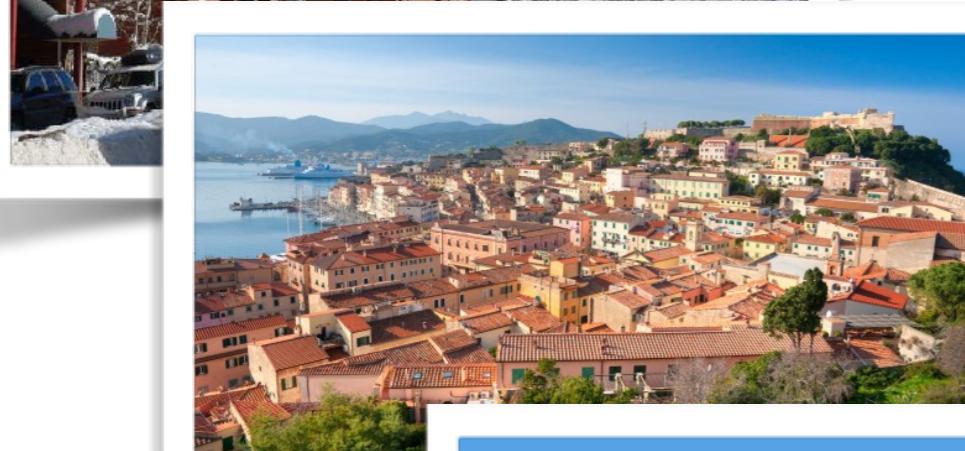


- Core Software Blocks in Quantum Chemistry, Asilomar, California, May 7-10, 2017 (**Organizer**: Anna Krylov)



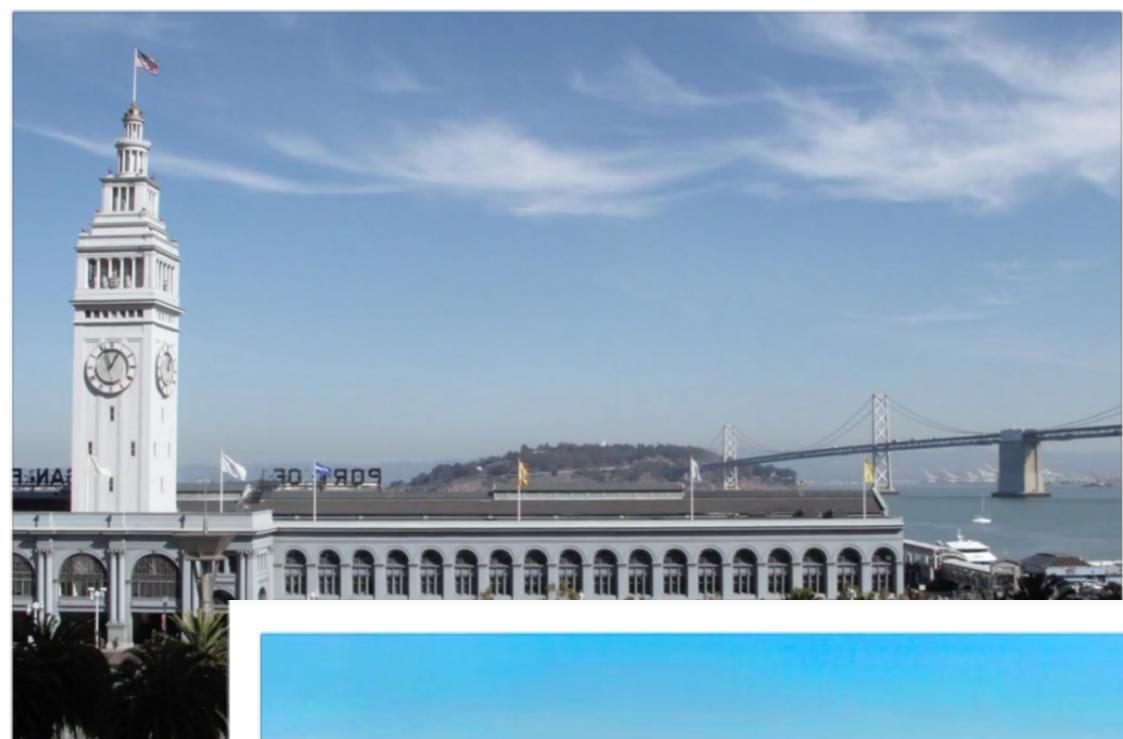
- Code and Data Interoperability, Virginia Tech, June 5-7, 2017 (**Organizer**: Theresa Windus)

MolSSI Workshops: Community-Driven



- Excited States: Electronic Structure and Dynamics, Telluride Science Research Center, July 17-21, 2017
(Organizers: Christine Isborn, Neepa Maitra, Xaosong Li, André Schleife)
- Coding Solvation, Livorno, Italy, August 23-25, 2017 (**Organizers:** Heather Kulik, Marco Caricato, Luca Frediani, Roberto di Remigio, Benedetta Mennucci, Oliviero Andreussi, Jógván Magnus Haugaard Olsen, and Filippo Lipparini)
- Best Practices in Molecular Simulation, NIST, Gaithersburg, Maryland, August 24-25, 2017
(Organizers: Michael Shirts and David Mobley)

MolSSI Workshops: Community-Driven



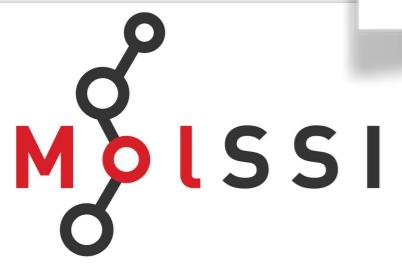
- Biomolecular Workflows, San Francisco, California, September 12-13, 2017 (**Organizers**: Daniel Smith, Aaron Virshup, Merry Wang, John Chodera)
- QM Schema, Lawrence Berkeley National Labs, November 30-December 1, 2017 (**Organizers**: Daniel Smith, Bert De Jong, Marcus Hanwell, Aaron Virshup)

MolSSI Workshops: Community-Driven



- Science Gateways and Computational Chemistry, New Orleans, Louisiana, March 2018, (**Organizers**: Nancy Wilkins-Diehr, Marlon Pierce, Katherine Lawrence, Sudhakar Pamidighantam)
- Python for Quantum Chemistry and Materials Simulation Software, Pasadena, California, June 2018 (**Organizer**: Qiming Sun)
- Tinker Modeling Software Workshop, Austin, Texas, June 2018 (**Organizers**: Pengyu Ren and Jay Ponder)

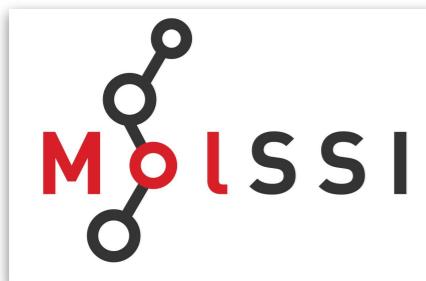
MolSSI Workshops: Community-Driven



- Modular Software Infrastructure for Excited State Dynamics, University at Buffalo, New York, June 2018, (**Organizer**: Alexey Akimov)
- Solving or Circumventing Eigenvalue Problems in Electronic Structure Theory, MolSSI HQ, Richmond, Virginia, August 2018 (**Organizer**: Volker Blum, Jianfeng Lu, Alvaro Vazquez-Mayagoitia, Chao Yang)
- Parallel Computing in Molecular Science, Berkeley, California, August 2018 (**Organizers**: Bert de Jong, Edward Valeev, Carlos Simmerling, Robert Harrison)

The Molecular Sciences Consortium (MSC)

- The Mission of the Molecular Sciences Consortium (MSC) is to develop a forum for the creation of a robust software development community within the molecular sciences. In particular, the initial scope of the MSC will be to:
 - establish standards in software and data;
 - lead in the development and evolution of licenses for software and data;
 - coordinate information exchange between stakeholders sharing/seeking data or developing particular software components, including SSE/SSI projects and community code.



Current Working Groups

- Forcefield Interoperability:

- Eliseo Marin-Rimoldi (lead), Chris Bayly, John Chodera, Erick Lindahl, David Mobley, Michael Shirts

- Data Formats for Quantum Mechanics Information (QM Schema):

- Daniel Smith (lead), Wibe de Jong, Marcus Hanwell, Paul Ayers, Geoff Hutchinson, Bob Hanson, Lori Burns, Matthew Chan, Sarom Leang, Joseph Montoya, David Sherrill, JR Schmidt, Benjamin Pritchard, Eduard Valeev, William Polik, Toon Verstraelen, Vamsee Voora, Julia Rice, Qiminig Sun, Ilya Kaliman, Eric Bylaska, Eric Berquist, Karol Langner, Theresa Windus

- Shared Memory Tensor Library Interfaces:

- Devin Matthews (lead), Theresa Windus, Evgeny Epifanovsky, Robert Harrison, Dmitry Lyakh, Saday Sadayappan, Beverly Sanders, Daniel Smith, Eduard Valeev, Anima Anandkumar, Anna Krylov, Beverly Sanders, Casey Battaglino, Grey Ballard, Azzam Haidar, Jiajia Li, Miles Stoudenmire, Paolo Bientinesi, Paul Springer, Ryan Richard, Shaden Smith, Stanmire Tomov

Proposed working groups



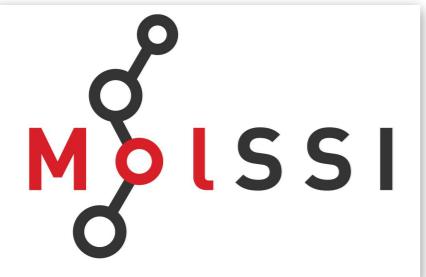
- Workflows
- Basis set description/tools
- QM/MM interfaces
- Software best practices
- Density-functional validation

Discussions at the MolSSI interoperability workshop, June 2017

MolSSI Headquarters @ Virginia Tech

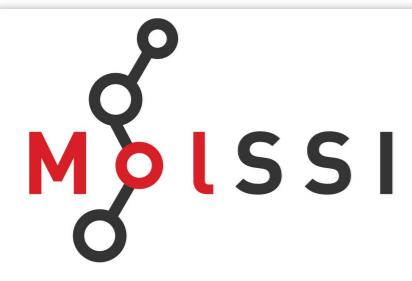


MolSSI occupies a newly renovated, 6,900 sq. ft. facility adjacent to campus.



MolSSI: Some Highlights So Far

- Hired **seven Software Scientists** in the last twelve months. Currently recruiting five more.
- **Ten software workshops** held so far; another six planned.
- New software components currently under development including **forcefield interoperability** frameworks, a **reference integral** implementation, a **new basis set exchange**, a **molecular simulation framework**, an **open QM database**, and more.
- **Community-driven working groups** established in forcefield interoperability, quantum chemistry data exchange, and tensor algebra interfaces.
- Held one software summer school and a bootcamp; another four educational workshops coming in 2018.
- **18 Software Fellows** now funded, and a competition for a third cohort ongoing.
- Proposal window for more **community-led Software Workshops** will open this summer.



Acknowledgments

- The many dozens of members of the CMS community who helped to develop the vision for the Institute over the last five years;
- NSF ACI-1547580.

Watch molssi.org for the latest information!