THE DOE CSGF ANNUAL

DEIXIS

2018

DEPARTMENT OF ENERGY COMPUTATIONAL SCIENCE GRADUATE FELLOWSHIP

BEATING A HEART DEFECT

Fellow Hannah De Jong's research could have immediate ramifications for people at risk of a cardiac condition.

PAGE 10

MORE FELLOWS' RESEARCH

- Jerry Wang tracks infinitesimal fluid flow
- Alnur Ali teaches computers a few things
- Jay Stotsky watches some unfamiliar films

ALSO: Alumnus Asegun Henry deejays some unusual beats, the Howes award recognizes two alumni, machine learning takes on fusion and a picnic plumbs the nanoscale world.

Announcing the James Corones Award

The James Corones Award in Leadership, Community Building and Communication recognizes the impact of mid-career scientists and engineers on their chosen fields across a range of areas.

Its namesake, Jim Corones, was a distinguished researcher and administrator who founded the Krell Institute, a nonprofit organization dedicated to serving the science and education communities. Under his guidance, Krell grew to supervise many projects and programs, most notably two prestigious Department of Energy-sponsored education initiatives: The Computational Science Graduate Fellowship (DOE CSGF) and the National Nuclear Security Administration Stewardship Science Graduate Fellowship (DOE NNSA SSGF). Jim retired from the company in December 2016 and died on April 28, 2017, after a long illness.

Broad eligibility: Mid-career researchers at a national laboratory or academic institution or in private industry.

Prize: A cash award of $2,000 and an engraved gift, plus travel for the winner to speak at one or two designated events.

For nomination procedures, deadlines and more information, including how to donate to the award fund, please visit https://www.krellinst.org/about-krell/corones-award.

INCOMING DOE CSGF CLASS

The newest class of Department of Energy Computational Science Graduate Fellowship (DOE CSGF) recipients – the 28th in the program's history – comes on board this fall. It's the first to include students pursuing a track in applied mathematics, statistics or computer science. All students receive yearly stipends, full tuition and fees and other benefits for up to four years.

Christiane Adcock
Stanford University
Computational and Mathematical Engineering

Sydney Andrews
Stony Brook University
Astrophysics

Kaley Brauer
Massachusetts Institute of Technology
Astrophysics

Jacob Bringewatt
University of Maryland, College Park
Physics

Kimberly Cushman
Yale University
Physics

Justin Finkel
University of Chicago
Computational and Applied Mathematics

Ryder Fox
University of Miami
Meteorology and Physical Oceanography

Steven Fromm
Michigan State University
Physics

Sarah Greer
Massachusetts Institute of Technology
Mathematics and Computational Science

Olivia Hull
Kansas State University
Physical Chemistry

Edward Hutter
University of Illinois at Urbana-Champaign
Computer Science

Dipti Jasrasaria
University of California, Berkeley
Chemistry (Physical-Theory)

K. Grace Johnson
Stanford University
Chemical Physics

Logan Kunka
Texas A&M University
Aerospace Engineering

William Moses
Massachusetts Institute of Technology
Computer Science

Samuel Olivier
University of California, Berkeley
Nuclear Engineering

Melissa Queen
University of Washington
Information Theory

Jesse Rodriguez
Stanford University
Plasma Physics

Lawrence Roy
Oregon State University
Computer Graphics

Noora Siddiqui
University of California, Irvine
Pharmaceutical Sciences

Steven Stetzler
University of Washington
Astronomy

James Sullivan
University of California, Berkeley
Astrophysics

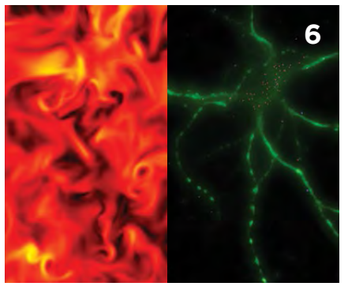
Anda Trifan
University of Illinois at Urbana-Champaign
Theoretical and Computational Biophysics

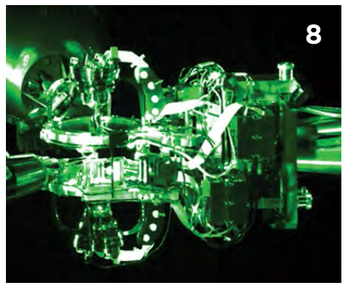
Michael Tucker
University of Hawaii
Astronomy

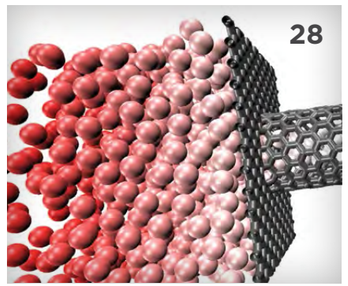
Caitlin Whitter
Purdue University
Computer Science

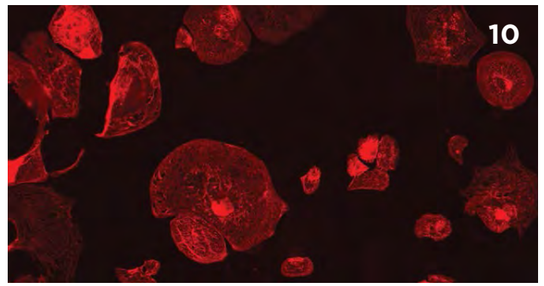
Paul Zhang
Massachusetts Institute of Technology
Geometric Data Processing

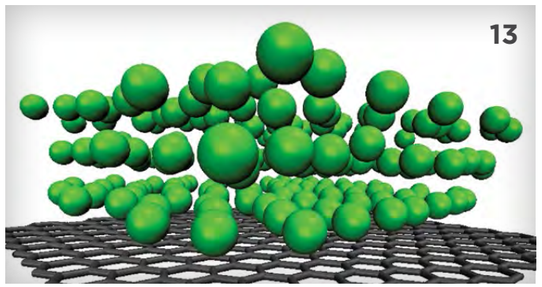
IN THIS ISSUE





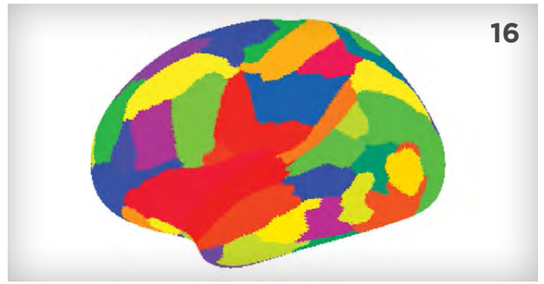


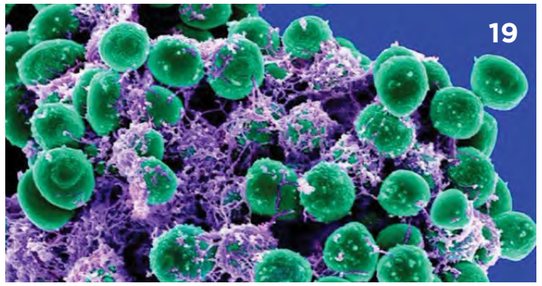




Departments

Features





DEIXIS, The DOE CSGF Annual is published by the Krell Institute. Krell administers the Department of Energy Computational Science Graduate Fellowship (DOE CSGF) program for the DOE under grant DE-FG02-97ER25308.

For additional information about the DOE CSGF program, the Krell Institute or topics covered in this publication, please go to: www.krellinst.org/csgf

Or contact:
Editor, DEIXIS
Krell Institute
1609 Golden Aspen Drive, Suite 101
Ames, IA 50010
(515) 956-3696

DEIXIS (ΔΕΙΞΙΣ — *pronounced dāksis) transliterated from classical Greek into the Roman alphabet, means a display, mode or process of proof; the process of showing, proving or demonstrating. DEIXIS can also refer to the workings of an individual's keen intellect, or to the means by which such individuals, e.g. DOE CSGF fellows, are identified.*

DEIXIS is an annual publication of the Department of Energy Computational Science Graduate Fellowship program that highlights the work of fellows and alumni.

DOE CSGF funding is provided by the DOE Office of Advanced Scientific Computing Research (ASCR), within the Office of Science, and the Advanced Simulation and Computing (ASC) program within the National Nuclear Security Administration.

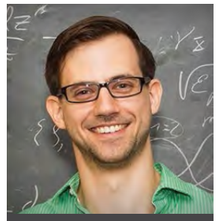
Manager, Science Media — Bill Cannon
Science Media Editor — Thomas R. O'Donnell
Senior Science Writer — Sarah Webb
Design — Stilt Studio, Inc.

ON THE COVER: Fellow Hannah De Jong creates DNA in the lab and algorithms on powerful computers as she seeks the causes of a damaging heart condition. Read about her Stanford University research starting on page 10. *Credit: Paul Sakuma Photography.*

DIVERGENT PATHS, SHARED PURPOSE

The award recognizes Davidovits and Middleton for outreach and service.

By Sarah Webb



Seth Davidovits

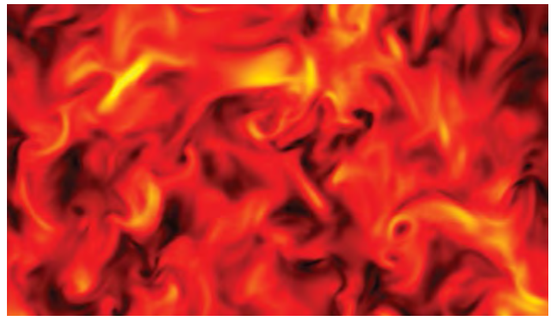
Sarah Middleton

At first glance, theoretical physicist Seth Davidovits of the Princeton Plasma Physics Laboratory (PPPL) and computational biologist Sarah Middleton of GlaxoSmithKline have little in common, with broadly different research interests and career paths. But the 2018 recipients of the Frederick A. Howes Scholar in Computational Science award share an enthusiasm for teaching, mentoring and bringing science into their communities.

As a child, Davidovits tinkered with the BASIC programming language, and he can't remember a time when he wasn't interested in science. As a Columbia University undergraduate, he was impressed with the directness of modeling physical systems via computers. That led Davidovits to combine physics and computation as a doctoral student and Department of Energy Computational Science Graduate Fellowship (DOE CSGF) recipient at Princeton University, where he developed new techniques to explore turbulent flow in plasma as it's compressed – processes that occur in fusion experiments, in the generation of X-rays and in astrophysics. In particular, he described – through theory and computation – the interplay between turbulence and heat in these systems. Simulations he's helped develop have shown that rapidly compressing flowing plasma quickly releases its turbulence as thermal energy, a phenomenon called sudden viscous dissipation. These results suggest ways to mitigate turbulence and harness it to boost efficiency in inertial confinement fusion and z-pinch experiments.

Since graduating in 2017, Davidovits has continued his work at PPPL under a DOE Fusion Energy Postdoctoral Fellowship. He has expanded his research to examine more turbulent systems and to study the effects of compression in two dimensions rather than three, simulations that are more relevant for z-pinch experiments. He's aiming for a career in academia or at a national laboratory.

Middleton's interest in science and computing didn't take off until she was an undergraduate at The College of New Jersey. She quickly discovered genetics and neuroscience but didn't feel at home in a laboratory. Instead, she developed an interest in programming, leading to a double major in biology and computer science and a future in computational biology research. As a Ph.D. student and DOE CSGF recipient at the University of Pennsylvania, she examined how the folding and location of RNA – the cell's genetic-information translator – in brain cells seeds learning and memory. Middleton still got her hands wet in the lab, examining individual mouse neurons and isolating and sequencing RNA molecules within. To analyze the resulting giant data sets, she applied her computational skills and created (with advisor Junhyong Kim) the program NoFold, which rapidly locates and groups complex RNA patterns that could represent cellular signals.



At the Princeton Plasma Physics Laboratory, Seth Davidovits explores turbulent flow in plasma as it's compressed. In their simulations, Davidovits and his colleagues discovered that rapidly squeezing a turbulent plasma can trigger the release of flow energy as heat. The results could be important for mitigating turbulence or boosting fusion experiments' efficiency. *Credit: Seth Davidovits.*

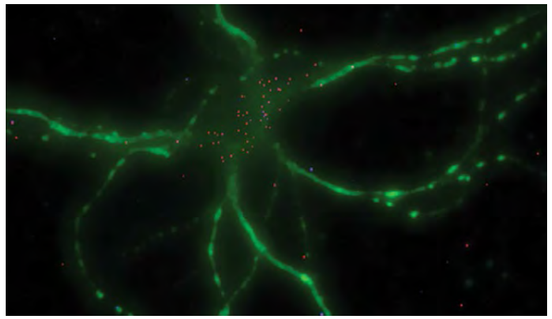
Middleton took an industry position at GlaxoSmithKline after graduating in 2017. She now works on a range of drug-discovery problems, from identifying genes that cause diseases to classifying disease subtypes. (Middleton is expecting a baby this summer and will deliver her Howes award talk at the 2019 DOE CSGF Program Review.)

As a doctoral student, Middleton recognized that her programming expertise was unusual – and desired – among biologists. She also remembered how overwhelmed she'd felt during her first computer science courses. "It seemed like everyone else had been programming since they were kids," she says. "That motivated me to see if I could help with that transition." She created an eight-session boot camp for biologists, aimed at scientists with zero programming knowledge and experience.

The workshop was wildly successful. For each of the three years Middleton directed it, she had to turn away students. Penn's Institute for Biomedical Informatics now offers a version of the course annually. During graduate school, Middleton also helped to create an online computational biology and genomics curriculum for high school students.

Both honorees have participated in events that bring science to the community. For several years Middleton organized activities for her department's booth at the Philadelphia Science Festival. She and her colleagues created bracelets that translated kids' names into DNA, teaching them about the genetic code.

Davidovits, meanwhile, has demonstrated physics concepts at the Plasma Science Expo during the annual American Physical Society Division of Plasma Physics meeting, sparking kids' interests with a Van de Graaff generator or the expansion and collapse of marshmallows under vacuum.



While a graduate student, Sarah Middleton predicted that RNA molecules (red, bluish and purple dots) localize to the different parts of a neuron (green, seen here in a mouse), including their dendrites (the spindly appendages). These RNA molecules could support an initial step in learning and memory. *Credit: Jean Rosario (Kim Lab, University of Pennsylvania).*

Mentors helped set Davidovits on his research path, and that's spurred him to work with high school and undergraduate students in the laboratory. This summer he'll supervise a Princeton undergraduate's project examining plasma for mass separation as a potential nuclear-waste remediation strategy. He also tutors high school students through Princeton's Community House After School Enrichment program, which supports underserved youth. He helps with homework problems and educational activities and has assisted students as they navigated college selection and financial aid.

Both Davidovits and Middleton credit the DOE CSGF as critical to their research success.

"That funding is game-changing," Davidovits says. It led him to explore options and find a productive research direction, and the annual review – the fellows' research meeting – broadened his knowledge of both computational methods and application areas.

Middleton agrees: "I met so many people through the fellowship and the meetings – really inspiring people." She says the fellowship gave her the "freedom to put forward the idea and far more leeway in how to pursue it."

ABOUT FRED HOWES

The Frederick A. Howes Scholar in Computational Science award, first presented in 2001, has come to stand for research excellence and outstanding leadership. It's a fitting tribute to Howes, who was known for his scholarship, intelligence and humor.

Howes earned his bachelor's and doctoral degrees in mathematics at the University of Southern California. He held teaching posts at the universities of Wisconsin and Minnesota before joining the faculty of the University of California, Davis, in 1979. Ten years later Howes served a two-year rotation with the National Science Foundation's Division of Mathematical Sciences. He joined DOE in 1991 and advocated for the fellowship and for computational science as manager of the Applied Mathematical Sciences Program.

Howes died unexpectedly in 1999 at age 51. Colleagues formed an informal committee to honor him and chose the DOE CSGF as the vehicle. With donations, including a generous contribution from Howes' family, they endowed an award in his name.

QUICK LEARNER

Katherine "Katie" Lewis is leader for the Applications, Simulations and Quality (ASQ) Division at Lawrence Livermore National Laboratory (LLNL). She also co-leads an effort to evaluate emerging hardware for processing neural networks and leads the Cognitive Simulation Project to incorporate machine-learning techniques into HPC simulations. She was an invited speaker at the 2018 DOE CSGF Annual Program Review.

DEIXIS: HOW DO YOU DEFINE MACHINE LEARNING?

Katie Lewis: It's a form of artificial intelligence that allows the computer to learn. Using lots of data and statistical methods, it can recognize patterns, defining probabilities based on the data that it has. Therefore, the computer can figure out a good guess about future behavior.

WHAT FACTORS LED TO USING MACHINE LEARNING IN PHYSICS?

In industry people have used machine learning to solve problems that we thought were unsolvable. Now image recognition from computing in some cases is better than human image recognition. That is astonishing.

HOW ARE YOU INCORPORATING MACHINE LEARNING INTO YOUR WORK AT LLNL?

We like to think of machine learning as a resource multiplier, making our simulations run smarter and faster. This strategy can help with tasks such as mesh partitioning, optimizing the way we divide problems to run on parallel processors to both balance the load and minimize the communication costs.

In addition, many simulations incorporate mesh movement. Turbulent behavior or other factors can cause the mesh to tangle and the simulations to crash. Then a user must step in and unravel the mess. With laboratory-directed research and development funding, we are examining whether machine learning can recognize tangling before it occurs or areas where the mesh should be allowed to move more while retaining the important physics that we care about.

HOW IS MACHINE LEARNING SHAPING OUR UNDERSTANDING OF PHYSICS?

In one project we have used machine learning to find unexpected results that could lead to more robust experimental conditions for studying thermonuclear fusion. For these inertial confinement fusion (ICF) experiments at LLNL's National Ignition Facility (NIF), researchers focus an array of lasers to compress and heat a tiny capsule of reaction fuel.

Physicists want to boost overall neutron yields in ICF experiments. But they also want to buffer the system against small perturbations that they can't control. So the team ran 60,000 simulations of ICF implosions and used them to train a machine-learning model. Running full simulations takes a lot of time, so this model offered an ideal system for varying parameters and getting a quick answer.

Based on conventional physics wisdom, the team had assumed that the ICF implosions must be shaped like a sphere. But the machine-learning model showed a novel solution – an ovoid, or egg-shaped, implosion – that they might never have considered otherwise. Then they ran a full simulation and got the same results. Since then the team has analyzed the simulation and come up with new physics theory to explain the unexpected results. The simulation still needs to be validated with experiments, but it's an example of how machine learning can help physicists examine a problem in a new way.

WHAT CHALLENGES DO THESE NEW APPROACHES PRESENT TO PHYSICISTS?

For a lot of the machine-learning methodologies the model results are black boxes. Researchers can't easily understand and interpret where they came from. That's not a big deal for search engines, but for our work we must evaluate results rigorously and understand the errors that we're introducing for every step.

We're also working to balance and weight experimental and simulation data in our machine-learning models. Machine-learning tools require huge amounts of data for training, and there will never be enough experimental data alone to create data-hungry deep neural networks.

Simulations are necessary approximations that complement experiments, but they introduce error. Models trained on simulations will inherit those problems. The question is how much that matters, but we can't expect them to perfectly match what an experiment does.

HOW IS MACHINE LEARNING SHAPING THE DEVELOPMENT OF NEW HARDWARE?

Just as video games sped the development of fast, affordable graphics processing units, or GPUs, that are now used in supercomputers, we expect that machine learning will spur hardware innovations.

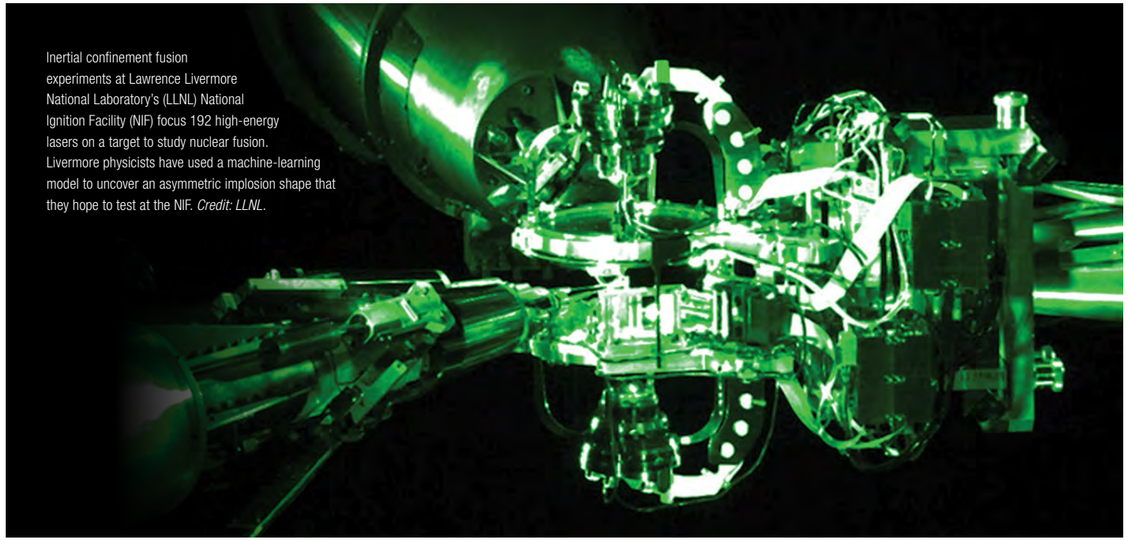
Neuromorphic computing is built on processors inspired by neurons in the central nervous system. Neurons collect synaptic signals – excitatory and inhibitory – from other connected neurons. Once the neuron's cell body reaches a specific voltage threshold, it spikes, sending a signal down its axon, which then connects to other neurons. IBM's TrueNorth chips, for example, have a spiking architecture, which mimics this behavior. Instead of using traditional if/then statements, such hardware uses the concepts of neurons, axons and synaptic weights to solve problems.

A neuromorphic processor can consume just one ten-thousandth of the energy used by a single GPU to solve similar problems. Therefore, matching these processors with the right problems could decrease computing costs and facility complexity.

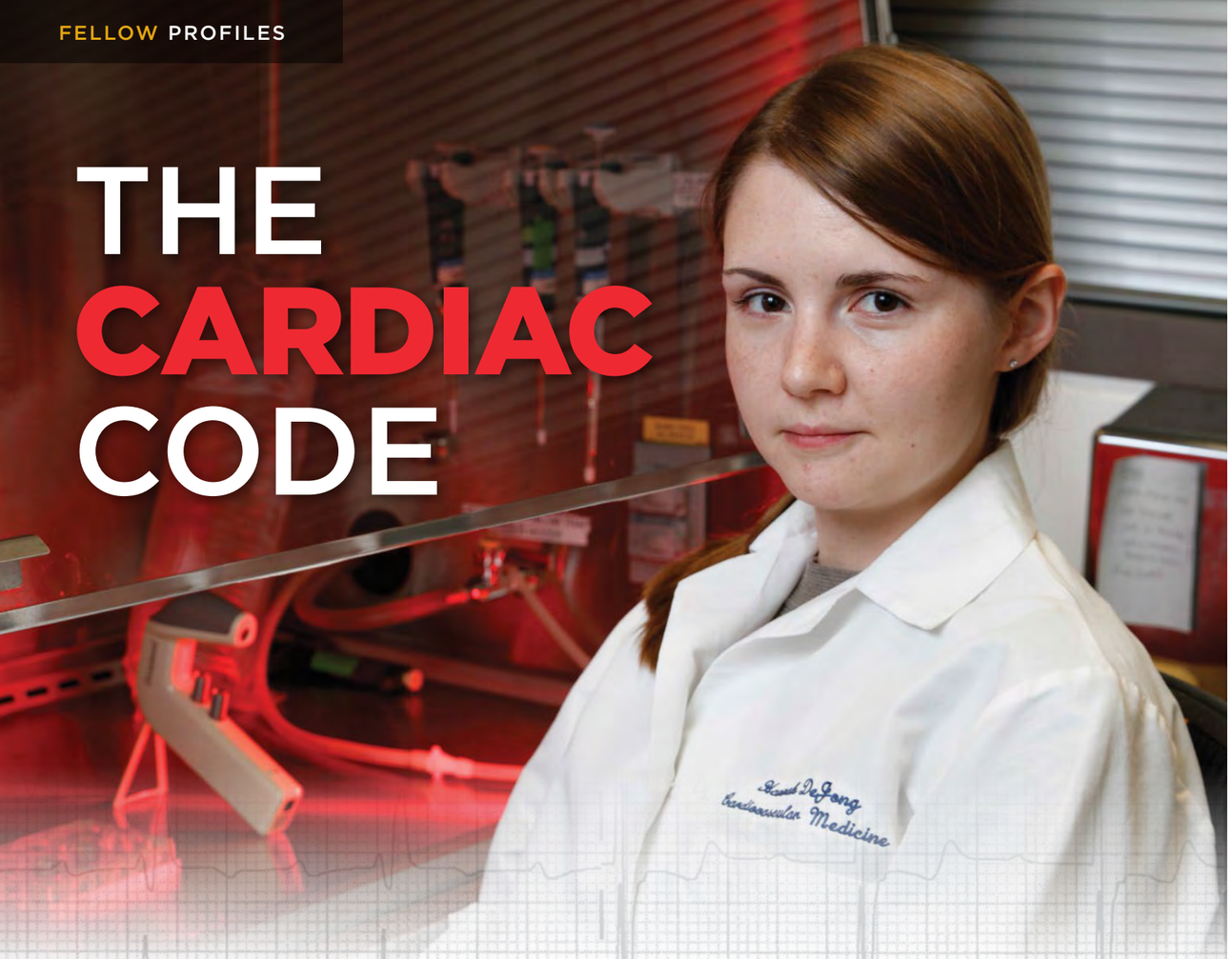
WHAT CAN NATIONAL LAB SCIENTISTS CONTRIBUTE TO THE FIELD?

We can leverage the work that industry has done in image recognition and classification and challenge those solutions with different data sets and problems.

We're not throwing away physics. We're looking at this as a hybrid that combines the fundamental physics with the probability-based methodologies of machine learning to get much better results.



Inertial confinement fusion experiments at Lawrence Livermore National Laboratory's (LLNL) National Ignition Facility (NIF) focus 192 high-energy lasers on a target to study nuclear fusion. Livermore physicists have used a machine-learning model to uncover an asymmetric implosion shape that they hope to test at the NIF. *Credit: LLNL.*



THE CARDIAC CODE

Combining experiments and computational analysis, Hannah De Jong seeks genetic signs of a damaging heart condition.

By Thomas R. O'Donnell

Hannah De Jong's heritage is in plant breeding. Her grandfather worked with potatoes, as does her father, now a Cornell University professor. Her mother studies tomato genetics at Cornell.

De Jong considered following the family into plant biology and did undergraduate research in the subject. But "I was always curious about research that could have an impact on human health care," she says. When De Jong started her doctoral studies at Stanford University, she switched from plants to human genetics.

Yet when deciding what to research, "honestly, the biggest factor is, is the problem exciting? I'm always just pursuing problems that are interesting to me."

The project De Jong (pronounced "De Young"), a Department of Energy Computational Science Graduate Fellowship (DOE CSGF) recipient, now pursues does indeed excite her. She's working with Euan Ashley to address a damaging genetic heart condition.

Hypertrophic cardiomyopathy, or HCM, enlarges and thickens part of the heart muscle, making it work harder to pump blood. "It's one of the main causes of sudden death in young people," says Ashley, a medical doctor with a Ph.D. and a professor of cardiovascular medicine, genetics and biomedical data science. Some patients may require a transplant, while others can live near-normal lives, perhaps with an implanted defibrillator.

HCM affects around one in 500 people in the United States. Ashley says he and De Jong hope understanding the genetics "will allow us to take a more precision approach to their condition and their treatment."

De Jong concentrates on MYH7, a gene that encodes beta cardiac myosin, a protein that is part of the motor that makes heart cells contract. "We don't know for sure exactly what the connection is between problems with that protein and HCM, but we believe that when that contraction doesn't occur properly the heart cells respond by making more of that contractile structure," enlarging the organ, De Jong says.

Researchers know it takes only one MYH7 mutation – one altered DNA letter, called a single nucleotide polymorphism (SNP) – to alter beta cardiac myosin. The trick is finding which SNPs out of hundreds do it. De Jong is systematically analyzing many variations to learn which cause faults in beta cardiac myosin.

The project was impossible a decade ago, De Jong says. "This will sound a little overdramatic, but the dream of many geneticists is to be able to precisely control which mutations are present in a gene and at what time and how the gene expression gets turned on." With new technology, "we're now able to come close to doing that."

De Jong's research focuses on a section of MYH7's DNA that's thought to be especially important in HCM's pathology. That means she must create and examine around 600 mutations.

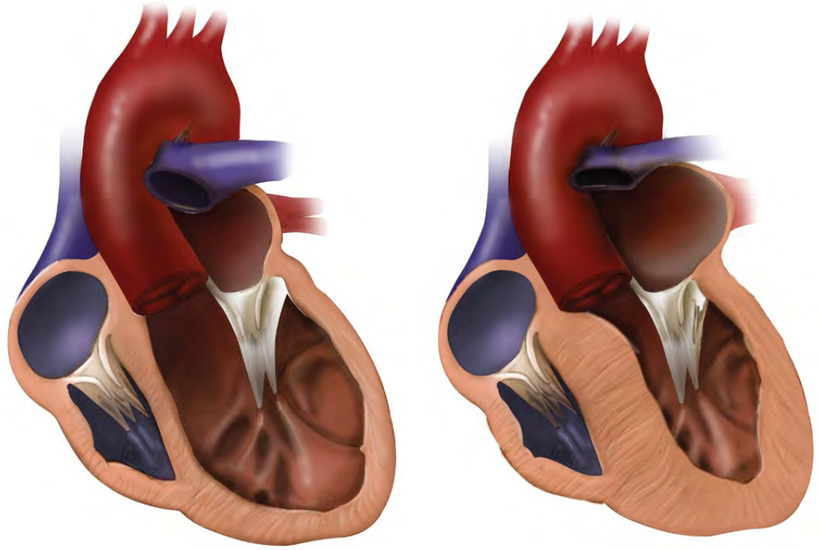
Her main tool is saturation mutagenesis, which creates multiple copies of pieces of the gene's DNA. Each contains a different SNP. De Jong uses computers to design chemical reagents that generate the mutations.

Once De Jong produces the mutations, she'll insert those not already seen in HCM patients into stem cells. After they grow into heart muscle cells, De Jong will sort them into three groups – small, large and in between – since those with HCM-associated mutations generally grow bigger.

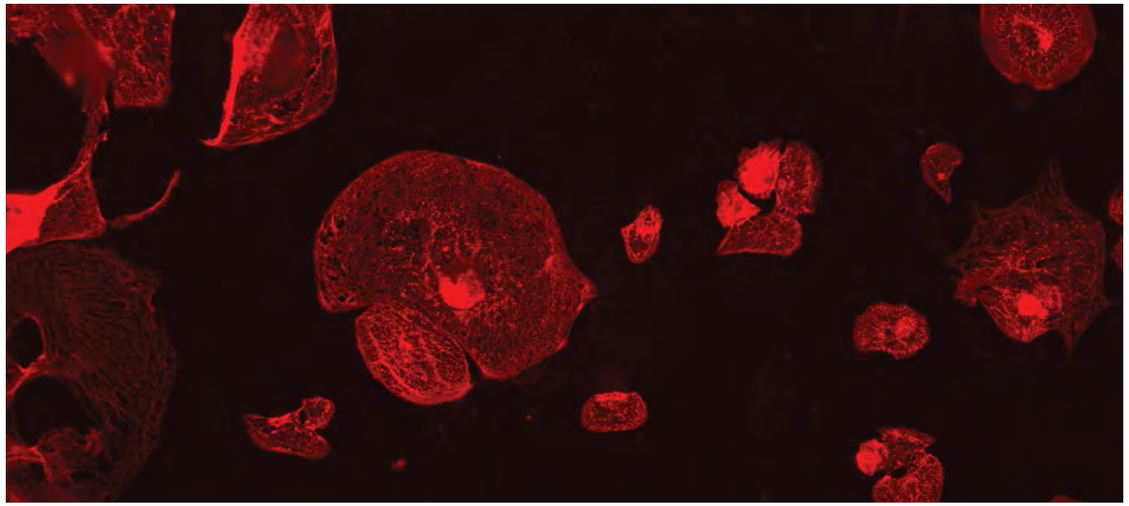
Finally, De Jong will decipher the cells' DNA and use computers to search the results for frequent mutations. Those that appear more in DNA sequences from enlarged cells probably are hypertrophic alterations that cause HCM.

There are many obstacles, especially inserting the DNA into stem cells. To simplify that step, De Jong is exploring another new technology: CRISPR/Cas9, which employs genetic machinery that bacteria use to defend their genomes from foreign DNA, such as that from a virus. CRISPR/Cas9 can be programmed to target specific DNA segments, letting scientists edit genes at specific locations.

"With CRISPR, you can make changes in the DNA of a living organism" rather than in a test tube, De Jong says. "You don't have to take out a piece of DNA, modify it and then attempt to put it back into a cell." But CRISPR can be inefficient and less specific when creating mutations.



Left: A heart with hypertrophic cardiomyopathy (HCM), on the right, showing muscle enlargement compared to a normal heart, on the left. *Credit: Megan Rojas.*
Opposite page, background: Electrocardiogram traces from multiple leads attached to an HCM patient's chest. Increased amplitude in QRS waves, most visible as large spikes on the V5 trace, are characteristic of this genetic condition. *Credit: Stanford Center for Inherited Cardiovascular Disease.*



This microscope image shows heart cells derived from stem cells and stained to show the contractile structure. *Credit: Hannah De Jong.*

With luck and skill, the researchers will have gigabytes of DNA sequencing data from cells exhibiting hypertrophic characteristics. De Jong then will spend less time in the wet lab and more at a computer. She relishes the change. "When I was younger, it was always a question of do I want to become a mathematician or a statistician, or do I want to become a (lab) scientist. I thought I had to choose." A high school bioinformatics internship helped her realize she could do both.

Analyzing her data, however, requires different computing capacity than the usual focus on maximum operations per second. "In genetics, generally the computational limitations are not processing-based, they're memory-based," De Jong says. She expects a specialized server in Ashley's lab will be sufficient to sift mutations from thousands of cells.

If it isn't, De Jong might get additional computer power through connections made during her 2016 Lawrence Livermore National Laboratory practicum. The lab's Catalyst system, a Cray CS300, is optimized for her research, with high memory capacity and a top speed of 150 trillion operations per second.

In the practicum, De Jong approached genetic engineering from another angle: detecting signs of DNA alterations rather than creating them. The group she worked with, headed by bioinformatics researcher Tom Slezak, studies bioterrorism. His team developed a system that analyzes air samples and identifies deadly viruses and bacteria.

But the technique can only spot previously known pathogens. Scientists fear that adversaries could genetically engineer microorganisms to make them more virulent or impossible to detect. So De Jong has developed software to identify artifacts that gene-engineering tools – particularly CRISPR/Cas9 – may leave in DNA sequences. The downside: There's no modified pathogen sequence data – thankfully – to test her technique. "As it stands, it's essentially my best guess."

Meanwhile, De Jong's Stanford research will inform scientists' fundamental understanding of beta cardiac myosin's nature, Ashley says, but also could be immediately relevant to patient treatment. When genetic testing finds rare mutations, counselors inform patients of the potential consequences. Those data can dictate, for example, "whether a family member who has that variant should be regularly screened or whether they're told that they're off the hook," he says.

For new or rarely seen variants, however, that part of the genetic test report is "a very short paragraph because there's so little we can say about it," Ashley adds. That means "the minute we have validated data from Hannah's experiment it will be relevant for the clinic" and genetic counseling.

Much work remains, though. De Jong's DOE CSGF appointment ends in 2018, but she expects to take an additional two years to finish her research. She's so dedicated to the work she took just two days off after her fall 2017 wedding.

Part of that urgency has come from meeting HCM patients. De Jong has also observed transplant surgery and examined hearts removed in those operations, donated to the lab with permission from the patients. It's driven home to her the human consequences of this devastating condition.

PEELING BACK LAYERS

MIT's Jerry Wang combines molecular simulations and physics to tease out how fluids organize and move in nanoscale spaces.

By Sarah Webb

In 2012, while spending a summer doing research at CERN, the famed particle physics laboratory in Switzerland, Gerald "Jerry" Wang had a rock-star science moment.

He'd never queued up for concert tickets or pulled an all-nighter for a Black Friday electronics deal. But that summer rumors buzzed around the institute that a groundbreaking announcement was imminent. When Wang heard that a special event was planned for the next day in CERN's auditorium, he and his friends came prepared. "We camped out overnight – sleeping bags, pillows, the whole nine yards," he recalls.

It paid off with prime seats for the announcement that physicists had experimentally detected the Higgs Boson, what Nobel laureate Leon Lederman coined as "the God particle."

"To see a community of thousands of people so uniformly excited about this thing that humanity's been working toward for 50-plus years – it was one of the distinct privileges of my life to be there."

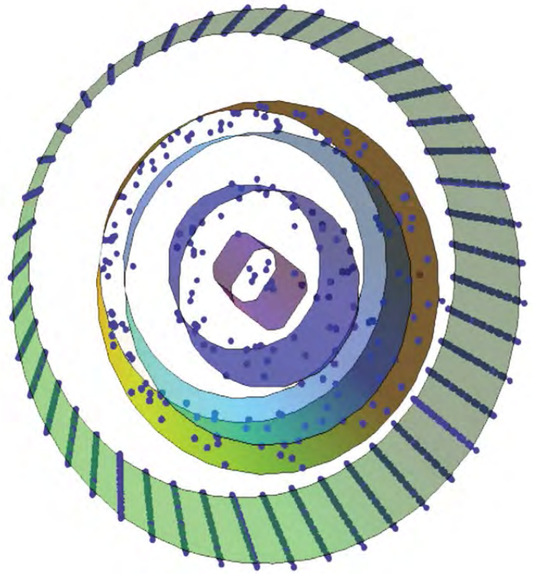
That enthusiasm has cemented the facets of Wang's science career. Today he's a Ph.D. student at the Massachusetts

Institute of Technology, using computational simulations to understand the properties of fluids at the nanoscale, where objects are thousands of times smaller than a hair's breadth. This fundamental understanding could be useful for a range of real-world applications, such as novel desalination technologies and fuel cells. A Department of Energy Computational Science Graduate Fellowship (DOE CSGF) supports his research.

As an undergraduate, Wang studied mathematics, physics and mechanical engineering. Besides spending two summers at CERN's Large Hadron Collider, where he worked with Yale University physics professor Sarah Demers, he pursued engineering research in fluid mechanics. The time at CERN "really opened my eyes to how you could answer such simple but unbelievably complicated questions using a ton of data and very powerful computers," he says.

At the time Wang applied to MIT in 2013, computational scientist and engineering professor Nicolas Hadjiconstantinou was filling a position in his group. Wang's skills in physics, molecular-level computations and statistical mechanics got him the post.

For Wang, it was a perfect opportunity to explore computational research and to split the difference between studying super-small, subatomic particles and macroscale fluids. As an engineer, studying the nanoscale behavior of fluids offered a chance to use his physics knowledge in ways that could have direct applications to real-world issues. The problem and computers are suited for each other, too. "Computers are so powerful now that we can actually simulate real systems that people are building at an atomistically resolved level," Wang says. "That's absolutely mind-blowing to me."



When confined within a carbon nanotube (CNT), fluid can adopt a layered structure in the form of concentric rings. This shows an equilibrium distribution of fluid within a CNT (points) along with theoretical predictions for fluid location (shaded surfaces). *Credit: Gerald J. Wang.*

For his Ph.D. research, Wang has been studying the structure of fluids confined to nanoscale spaces and how those trapped fluids transport energy.

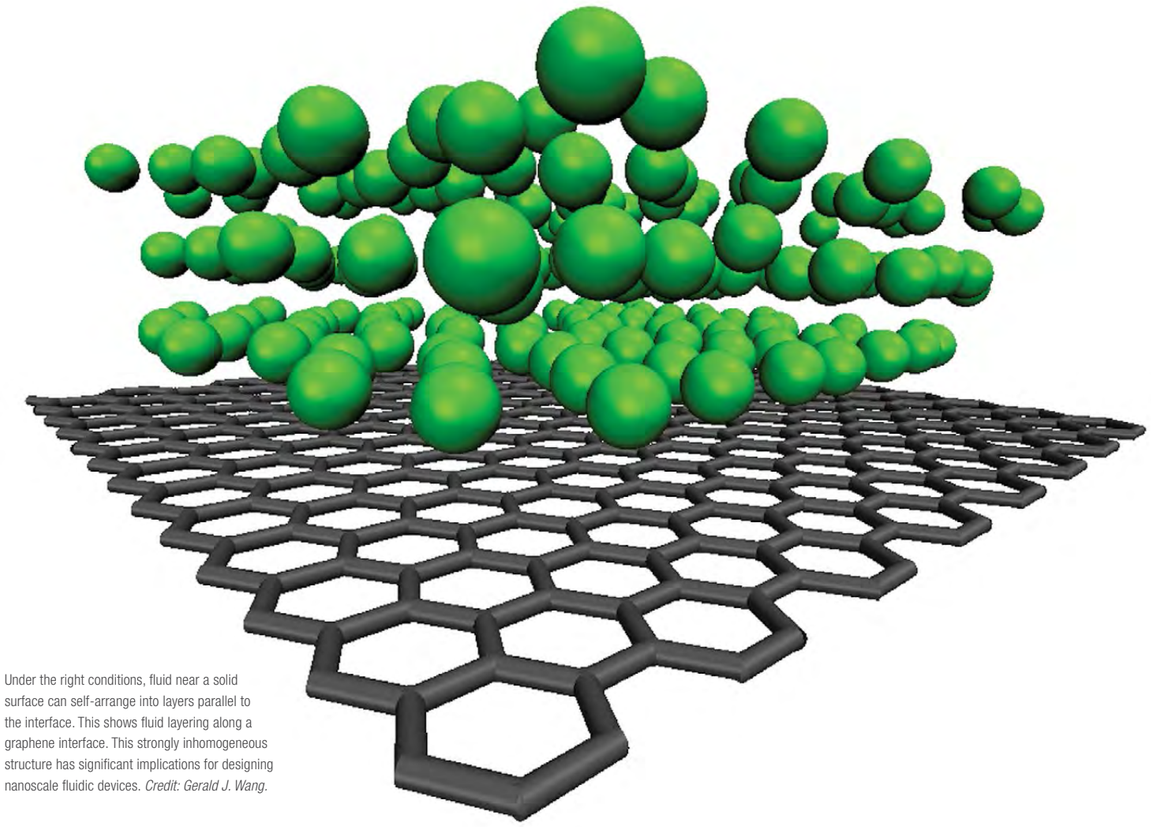
Under typical conditions, transferring energy from a solid to a fluid is relatively inefficient. On a hot day, Wang notes, you'll cool off much faster by pressing an ice bag against your face than by sitting in a room filled with cool air.

Since the 1990s, researchers have observed that fluids next to a solid can form layers parallel to the solid's surface. But few molecules in everyday fluids exhibit any particular pattern or structure. In super-tight spaces, though, such as capillaries just nanometers in diameter, "the molecules very often have little choice but to settle like a solid," Wang says. Bands of molecules form parallel to the surface confining them, like layers in a cake. Using computational simulations, Wang has outlined the theory for how these layers form and how far they extend. He's also examined how material density, temperature and interactions between molecules in the fluid and solid affect the fluid layers.

In these nanoscale channels, heat transfer between solids and liquids looks far different than it does at a larger scale. "Suddenly when you try to transfer heat from the solid to what you would imagine is the liquid, you start seeing heat transfer rates that are much, much higher than you would ever imagine is possible for a solid to a liquid," he says.

With his physics background, creativity and ability to build simulations, Wang has made significant contributions to understanding a 40-year-old problem, Hadjiconstantinou says. But Wang's ability to find connections and communicate them vividly, through images such as the cake analogy, make him a standout scientist. "Jerry is just very, very creative in everything he does."

Wang is especially excited about how these principles could be used in everyday life, noting that "access to clean water is one of the great challenges of our time." Cleansing dirty water and pulling salt from ocean water both require overcoming the same challenge: removing nanoscale impurities. Engineering such systems demands a detailed knowledge of the fundamental physics in play at that size. "That's your best shot at actually designing a system in a highly functional and intelligent way. So that's the kind of gap in knowledge that I try to work to fill."



Under the right conditions, fluid near a solid surface can self-arrange into layers parallel to the interface. This shows fluid layering along a graphene interface. This strongly inhomogeneous structure has significant implications for designing nanoscale fluidic devices. *Credit: Gerald J. Wang.*

Today's desalination strategies rely on reverse osmosis, which demands a large energy investment. If engineers can design membranes with pores that allow water molecules but not salt ions to pass, they could efficiently produce potable water from salt water. "There's no better way to tackle that than to use devices that are attuned to that length scale," Wang says.

By improving understanding of energy transport at the nanoscale, his research could also help engineers design better batteries with longer lives – an application that would benefit everyone from smartphone users to people in developing countries who lack reliable power.

For his summer 2015 practicum at Argonne National Laboratory, Wang modeled a different nanoscale energy transport phenomenon: waves known as inhomogeneous surface plasmon polaritons, which occur in metals such as gold and silver and in other, more exotic materials. Understanding these waves could be important for designing nanophotonic devices, minuscule light-interacting components that can be used in electronics, computer memory, solar cells and many other applications. Though very different from studying fluids, the project was an interesting transport problem at the nanoscale, Wang says, and a valuable introduction to computational science at the national laboratories and on DOE supercomputers.

Wang's Argonne advisor, Stephen Gray, was impressed with his motivation, enthusiasm and quick mastery of the research. "I feel very lucky that Jerry worked with me," Gray says. Though Wang was at Argonne for only a few months, the team is finishing up odds and ends on the project and plans to publish its results in the coming months.

As he finishes his doctorate, Wang hopes to continue his career in nanoscale computation and engineering. Whether he ultimately lands in academia or at a national laboratory, he looks forward to making molecular simulations useful for engineering. He wants to continue to bridge the gap between experiment and engineering design and the computations that inspire so much of it.

Gray expects that Wang's enthusiasm will take him far. Besides his technical talents, Wang also approaches challenging problems fearlessly, he says. "It makes him a really remarkable individual."

HIGHER LEARNING

Alnur Ali rides an early career at Microsoft to the upper ranks of machine-learning research.

By Monte Basgall

By the time Alnur Ali began his Department of Energy Computational Science Graduate Fellowship (DOE CSGF) in 2014, he'd already charted an odd path to Ph.D. studies at Carnegie Mellon University.

Inspired early on by a world-famous physicist's book, he focused on computer science and mathematics at the University of Southern California and took a summer's worth of physics at the University of Cambridge in England.

But he graduated with "no real way to combine those three areas of studies," Ali now recalls. Feeling divided, he spent 2004 to 2013 as a Microsoft software engineer, interrupting an academic quest for an industrial one but picking up new skills in the process.

At first he focused on "important problems in the field of information retrieval, like ranking and query rewriting," he says, to improve Microsoft's Bing search engine. But the experience also sowed seeds for a future in machine learning. This growing field uses special algorithms that guide computers to teach themselves to search widely for hidden but related additional information, sometimes via neural networks modeled after the brain. Besides providing Ali a well-paying job, Microsoft exposed him to "a ton of experts (who) showed me what it meant to be a machine-learning researcher early on, what kind of questions to ask and how to think about problems."

He explains that "machine learning is partly about making those models work in the real world, and partly about understanding from a mathematical point of view how those models behave. And it partially has origins in neuroscience." Because he was "interested in how the brain works, I was hooked right away."

Microsoft encouraged Ali to work with the Seattle research community, including the statistics faculty at the University of Washington. Though he calls those interactions "a very interesting, cool challenge," he quickly notes that "I always knew I wanted to go back to graduate school."

His decision bucks a trend. After such a significant time away from college, "it is less common for students to enter a Ph.D. program in computer science," says Zico Kolter, an assistant professor in the subject and a member of Carnegie Mellon's Machine Learning Department who became Ali's first Ph.D. advisor.

"I believe this is largely due to the fact that software engineers are in very high demand in the current job market. Alnur was indeed exceptional both in his drive to come back to grad school and in his ability to find a way to do research while being employed at a large company. My impression is he is drawn to understanding mathematical algorithms at a more fundamental level than one is typically exposed to in industry."

Ali enrolled at Carnegie Mellon as a doctoral candidate in machine learning in 2013, and his research took him to Stanford University during 2014.

"Machine learning is everywhere these days," he says. "So it's important to theoretically understand the pros and cons of different methods." Those include "their predictive accuracy, how long they take to run and how easy they are to use."

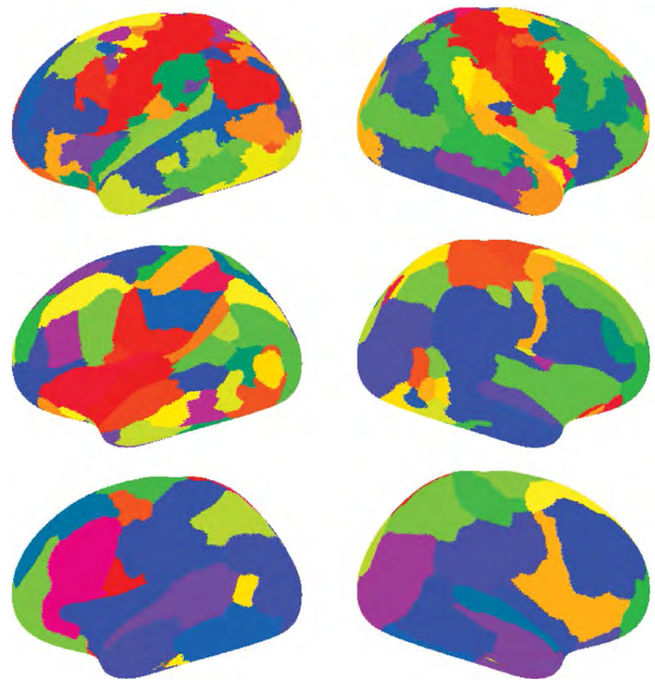
"In part, that boils down to working a bunch of math to explain these tradeoffs. The other part is applying machine learning in new ways that can help people, which sometimes requires developing faster algorithms. As I've gone through my Ph.D., I've wanted to do more good for society at large."

Two machine-learning methods Ali studied for his 2015 DOE CSGF practicum at DOE's Lawrence Berkeley National Laboratory have influenced his approach.

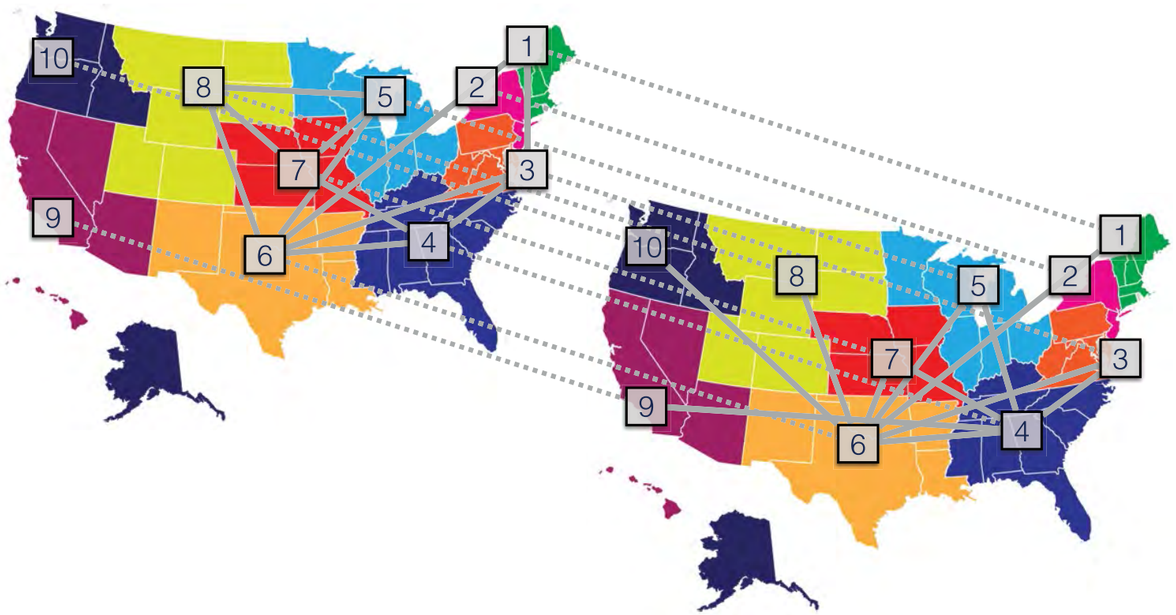
The first, the inverse covariance matrix, is "basically related to how correlated two objects are," he says. The second, pseudo-likelihood, is "better at estimating correlations when you have a lot of measurements, like in big data."

He employed both methods for a 2017 paper he co-wrote – with practicum advisor Sang-Yun Oh and two other collaborators – for the 20th International Conference on Artificial Intelligence and Statistics in Fort Lauderdale, Florida. That study focused on so-called high-dimensional, or big data, situations "where the number of variables is possibly much larger than the number of data samples," they wrote. They demonstrated how those methods apply to real-world finance and wind-power data.

With Kolter and Ryan Tibshirani – his second doctoral advisor – Ali also tested those techniques in a 2016 paper written for the 30th Conference on Neural Information Processing Systems in Barcelona. The authors used them to study multiyear spreads of influenza around the United States. Ali is now working with Oh on a third paper that uses those methods to characterize functional relationships between brain regions where a single pixel location within a functional magnetic resonance imaging brain scan is the smallest unit, says Oh, now an assistant professor of statistics and applied probability at the University of California, Santa Barbara.



Each row presents a segmentation of left and right hemispheres of the human cerebral cortex computed by different methods. Row one: the results of HP-Concord (with persistent homology fine-tuning), a technique Alnur Ali and colleagues have developed. The data-driven method's segmentations resemble those derived from neuroscience methods. In contrast, segmentations in the second and third rows, from other data-driven methods (middle, HP-Concord plus Louvain; bottom, thresholded sample covariance matrix plus Louvain), appear to lack fine detail. *Credit: After a figure in "Communication-Avoiding Optimization Methods for Distributed Massive-Scale Sparse Inverse Covariance Estimation." Penporn Koanantakool, Alnur Ali, Ariful Azad, Aydın Buluç, Dmitriy Morozov, Sang-Yun Oh, Leonid Oliker, and Katherine Yelick.* Proceedings of the 21st International Conference on Artificial Intelligence and Statistics (AISTATS), *2018.*



A model of influenza spread developed by Alnur Ali and colleagues. A solid line between two regions indicates that the numbers of flu reports in the regions are statistically dependent. A dotted line between any two regions (i.e., across the two maps) indicates that flu reports from the region in the upper left map can predict reports coming one week later from the adjacent region in the bottom right map. *Credit: After a figure in "The Multiple Quantile Graphical Model." Alnur Ali, J. Zico Kolter, and Ryan J. Tibshirani.* Advances in Neural Information Processing Systems 29 (NIPS), *2016. Based on data from the Centers for Disease Control and Prevention.*

For the ongoing brain study and the 2017 paper, Oh and Ali have relied on Edison, Berkeley Lab's massive, 2.57-petaflops supercomputer.

"Data-driven scientific discovery is an important trend in many research fields," Oh says. "Alnur's research interest and skill set is in the sweet spot of this trend." Ali, he adds, also is "a kind and considerate person with grounded optimism."

Ali says he's grateful for the fellowship's financial support and for the knowledge and experiences it has provided. It has made him "more interested in doing work on the theoretical side (without) losing sight of the real-world applications. Also, I am more willing to take risks with my research."

Tibshirani, a Carnegie Mellon associate professor of statistics who also focuses on machine learning, agrees that Ali is "passionate about doing research that has social value. He also has a way of speaking and describing his research that is easy for others outside his specialty to understand. That is something that is mature and hard to teach."

A Toronto native whose parents are from India, Ali went to high school in Singapore when his father was assigned there for business. That's where Ali's mathematics teacher suggested he read *Surely You're Joking, Mr. Feynman*, a semi-autobiography by Richard Feynman. The impish Nobel-winning theoretical physicist, who died in 1988, made major discoveries in quantum electrodynamics and superfluidity but also was conversant in safecracking and bongo drumming.

"That sparked something in me," Ali recalls. "(Feynman) was very mischievous and very curious, and I related to that on some level. From then on I got a lot more into math." Also, "one day my dad brought home a computer. At first I just played games on it. Then I started wondering how I could make my own game." That led Ali to try coding.

So, many years later, what will he do when he gets his Ph.D.?

"I'm pretty much open to any place where I can work on interesting problems with interesting people and make a difference in the world."

THE BIOFILM TANGO

It's only several seconds long, but Jay Stotsky's computation composes one of the world's most accurate models of biofilm movement.

By Jacob Berkowitz

As an undergraduate studying chemical and biological engineering at Tufts University, Jay Stotsky helped pay his way by working as a piano accompanist for aspiring vocalists, tuning his skills on a variety of musical styles, from Baroque to Broadway.

Though music is part of his life's rhythm, math and science have provided the driving bass beat.

"I've been interested in math for as long as I can remember," says Stotsky, a classically trained pianist and native of the Boston suburb of South Easton.

"There are a lot of similarities in the approach to how you learn a piece of music very well versus figuring out a mathematical problem. There's an attention to detail that's very important in both."

As a Department of Energy Computational Science Graduate Fellowship (DOE CSGF) recipient at the University of Colorado Boulder, Stotsky has applied his zest for attentive problem-solving to models of how colonies of bacteria, or biofilms, move when water flows over them.

The topic is as important as brushing your teeth or, in technical terms, the daily act of applying stress-strain to remove an oral biofilm.

Stotsky's University of Colorado advisor David Bortz began chewing on the problem as a postdoctoral researcher at the University of Michigan in the early 2000s.

As part of the mathematical biology group, Bortz teamed with John Younger, an emergency room physician now at Akadeum Life Sciences , and chemical engineering professor

Above: Scanning electron micrograph of a clump of *Staphylococcus epidermidis* bacteria (colored green) in an extracellular matrix, a key biofilm component that connects cells and tissue. *Credit: National Institute of Allergy and Infectious Diseases.*

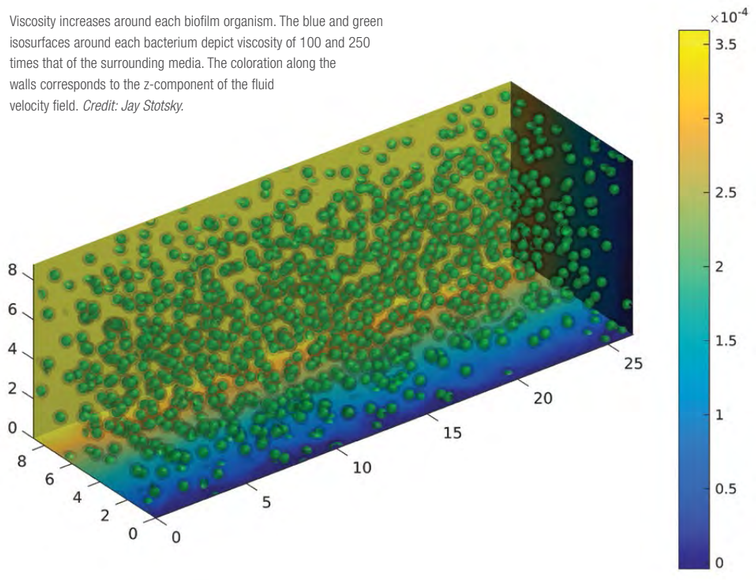
Michael J. Solomon to study the behavior of bacterial aggregates in the bloodstream. Soon he was researching a central problem in this field: the dynamics of potentially life-threatening biofilms growing in catheters and IV (intravenous therapy) lines.

Then, soon after Bortz arrived at the University of Colorado in 2006, he was introduced to researchers at the nearby National Institute of Standards and Technology (NIST) campus who were concerned about biofilm growth in biofuel pipelines.

“It was a similar problem in the two cases,” Bortz says. “You had a biofilm growing in a pipe, and you wanted to know the stress-strain relationships of the biofilm. How easy would it be to break it apart and get rid of it, and what could you do to the surface of the pipe that would make it harder to grow on?”

A biofilm consists of bacteria physically interconnected via a web of polysaccharides, or sugar chains – the extracellular matrix – to create a viscoelastic fluid with “the consistency of mucus,” Bortz says.

“It’s been a really challenging problem to think about because the biofilm viscosity changes by a factor of 500. It’s really viscous adjacent to the bacteria and then a few microns away it drops to the viscosity of water.”



Viscosity increases around each biofilm organism. The blue and green isosurfaces around each bacterium depict viscosity of 100 and 250 times that of the surrounding media. The coloration along the walls corresponds to the z-component of the fluid velocity field. *Credit: Jay Stotsky.*

To tackle the problem, Bortz adapted the Immersed Boundary Method, a technique first developed in cardiology to solve fluid-structure interaction questions.

His Michigan collaborators used a laser confocal microscope to identify the three-dimensional positions of the bacteria. Bortz used that information to create a prototype model that, for the first time in such simulations, treated the biofilm’s macroscale rheology (deformation and flow properties) as a feature that results from the bacterial interactions.

When his first Ph.D. student on the project graduated in 2012, Bortz canvassed the first-year graduate numerical analysis class for someone to pick up the baton on the heterogeneous rheology immersed boundary method, or hrIBM.

With his academic background, “Jay was the perfect fit for the project,” Bortz says.

The professor soon saw that his graduate student’s intelligence, creativity and experience were matched by his determination. “Jay has this amazing ability to dive down to the core of the problem and just not stop until he’s solved it.”

Stotsky began work to improve the hrIBM model’s computational performance in October 2013. Running it on Janus, the university’s supercomputer (since decommissioned and replaced by Summit), Stotsky applied highly efficient methods and computational techniques for solving the equations governing biofilm-fluid interactions. By early January he’d accelerated the simulations by an order of magnitude.

“In three months he’d taken the model to the next level in terms of speed and accuracy,” says Bortz, who enthusiastically sponsored Stotsky for the DOE CSGF.

“The big question at the time,” Stotsky says, “was how well does the model really compare to the experimental data.”

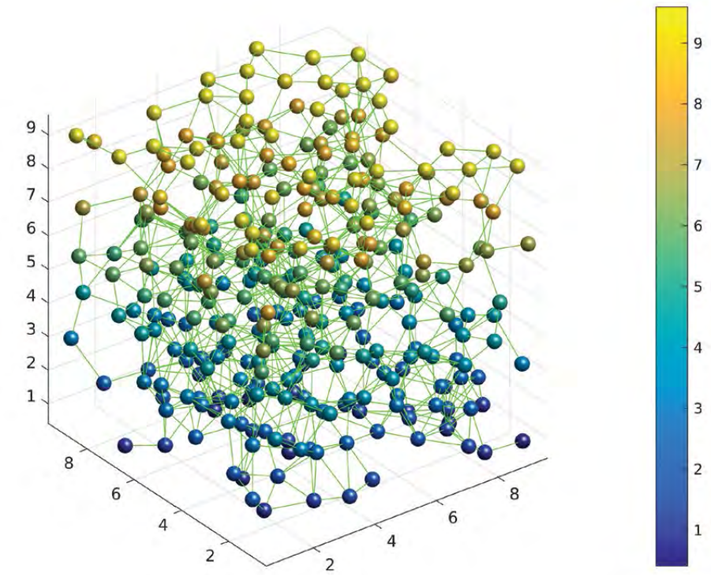
He set to work validating it using detailed experimental data from live biofilms of *Staphylococcus epidermidis*, a skin bacterium that turns deadly when it infects catheters. The University of Michigan collaborators grew the biofilm and tested its rheological properties.

Stotsky explored a half-dozen mathematical models for the stress-strain relationships and ran simulations that portrayed several seconds of 3-D biofilm movement, each one requiring a full day on Janus.

He discovered that when the spring-like connections between the bacteria were modeled in a particular way, nearly all of the computed data points hewed to the experimental data.

“Nobody before was able to hit the values of an experiment so precisely,” Bortz says.

As first author of a 2016 paper in the *Journal of Computational Physics*, Stotsky reported that “the hrIBM model is the first that can accurately compute bulk material properties of biofilms” based on coupling their microscopic connections with the extracellular matrix’s varied rheology.



Cell location in a virtual biofilm section. Each sphere represents a bacterium, and the green lines between cell pairs depict viscoelastic links between those bacteria. Sphere coloration corresponds to a cell’s distance from the biofilm’s bottom. *Credit: Jay Stotsky.*

Having validated the hrIBM model, Stotsky next set out to extend this work by developing a statistical model that links biofilm rheology to how the bacteria are distributed in space.

“With the statistical model, the idea is to be able to create artificial biofilms that have similar properties as a real biofilm would,” says Stotsky, who brought to the work insights in applied mathematics techniques gained during his 2016 DOE CSGF practicum at Lawrence Berkeley National Laboratory under Dan Martin and Phil Colella.

Stotsky developed a spatial statistics model using data sets of 4,000 or so bacterial positions as input. It characterizes fundamental statistical interactions between the organisms using various means of calculating the dependence of that interaction on their proximity.

After using the statistical model to simulate a biofilm colony, bacterium by bacterium, Stotsky paused to overlay his simulated data on the experimental data points. They overlapped with “close approximation.”

“I’d been working on it for several months and hadn’t had a chance to check and see if it was doing what I thought it was, so I was very happy to see that it was,” Stotsky says with a laugh.

The model revealed that the bacteria have a “favorite distance” by which they’re separated, and that in biofilms there’s strength in disorder.

“The most surprising thing is that I found that having just a little bit of non-uniformity, for example not having the bacteria aligned on a grid, tends to make the biofilms a little stronger. I thought it would be the other way around.”

The software for generating biofilm models will be freely available on the group’s website, Bortz says.

As Stotsky prepares to graduate in spring 2018, he’s applying for postdoc positions and still performing music, now in a tango group, leading dancers in their complex, sweeping moves.

“It’s very difficult dancing,” says Stotsky – exactly the kind of exceptionally detail-oriented interaction between music and movement, or math and movement, he will continue to produce.

PERIODIC TABLE PLAYLIST

Energy scientist Asegun Henry slows down atomic vibrations in elements and listens carefully.

By Bill Cannon

On the commute from his Georgia Institute of Technology lab, Asegun Henry rolls through Atlanta's congestion in a sonic bubble of hip-hop and R&B, neo soul and reggae. But it's a cinematic moment that most closely captures the sounds he models in his work as assistant professor at the George W. Woodruff School of Mechanical Engineering.

"Sonifying elements" – audio counterparts to visualizations, in this case what you might hear if you could hold your ear to the periodic table – sound like the synthesized effects in the Jodie Foster movie *Contact*, Henry says. It's the scene in which scientists fire up a machine designed to punch a wormhole in space to engage extraterrestrials. Like that device, fundamental elements hum rhythmically – imagine an industrial washing machine backed by a distant lawn mower. "My favorite is the first element we did: silicon."

Henry participated in the Department of Energy Computational Science Graduate Fellowship (DOE CSGF) from 2005 until 2009, when he earned a Ph.D. from the Massachusetts Institute of Technology. At first, he lamented the required computer science classes eating into his physics and engineering course load, but his regard for the fellowship has only grown over the years. "Just about every paper I've published since graduate school leverages this computing knowledge," he says. Everyone in his lab writes codes scalable to high-performance-computer simulations from the get-go, which confers a competitive edge. "Other people in my field are limited to commercial or open-source codes, and most don't know how to write code for large parallel machines."

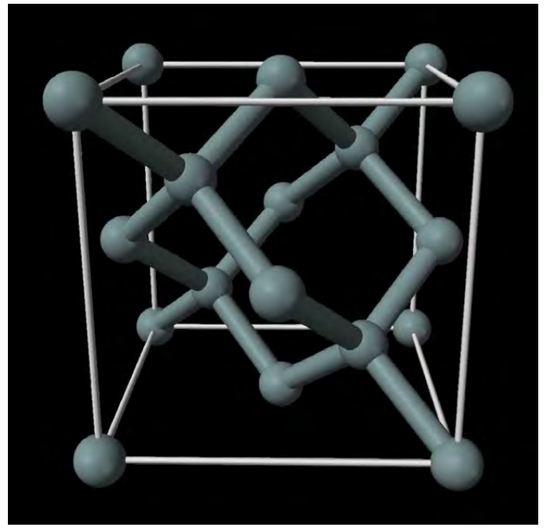
Programming is particularly valuable in his studies of molecular simulation and phonon transport, keys to problems in energy conversion, transport and storage. The sonification work also relies on coding to tease out molecular patterns in materials that may be advantageous in all of the above.

"We tend to do statistical analyses," which are ideal for describing how atoms behave on average, Henry says. "But in doing that, you throw away any particular pattern information." To sift for those patterns, Henry derives audio signals from the simulated movement of vibrating atoms in elements.

And ears, it turns out, are much better at sampling for patterns than eyes scanning the same data. For example, visualizations of Gladys Knight and the Pips' versus Marvin Gaye's "I Heard It Through the Grapevine" would bear no resemblance to one another. But your ear can instantly discern they're different versions of the same tune.

Atoms vibrate on a picosecond scale, in the terahertz, or trillions of hertz. "The human ear can resolve only 20 kilohertz,'" Henry says, or "about a billion times slower." That necessitates decelerating vibrations and retaining the amplitude data while re-labeling the time axis proportionally. For instance, if two parts of the signal differ by 1 picosecond, Henry dials it down to a millisecond, preserving the relative differences in a frequency's peaks and valleys. "Now that it's on the timescale of milliseconds it's kilohertz, and we can hear it."

Henry says the idea to sound out the elements came to him even before he started undergraduate studies at Florida A&M University. An audio signal originates as vibrations; since atoms vibrate in a solid, he reasoned, they must produce a signal, and "I always wondered what that audio signal would sound like. I just never had the tools (to investigate it)." In graduate school, he learned it was possible to simulate atom movements like vibrations but was "sidetracked with a bunch of other stuff."



A model of crystalline silicon from a frame of Asegun Henry's favorite atom-music video, available on YouTube: https://www.youtube.com/watch?v=vagPz9vf8dw. Henry generated the sound from a single atom's velocity versus time, slowed down to present 5 nanoseconds of vibration over 50 seconds so that it could be heard by a human ear. *Credit: Asegun Henry*

Ears, it turns out, are much better at sampling for patterns than eyes scanning the same data.

When he started at Georgia Tech in 2012, Henry revisited the subject. But he faced a new problem: "I didn't know how to write or create a sound file from scratch from a piece of data." So he did what any professor might do: He asked a summer undergraduate to investigate. The student found a one-line command in MATLAB, MathWorks' ubiquitous science and engineering software, called "wavwrite" that "can kick out a formatted WAV file of your oscillatory signal."

That small but important step helped lead to a National Science Foundation five-year award for early-career scientists to catalog the periodic table's musical signatures. It also made possible, in 2017, the first paper, in the *Journal of Physical Chemistry A*, describing a method for sounding out a scientific problem, in a molecule called polythiophene – a polymer of interest for its novel heat-conducting properties. His co-authors were Georgia Tech musicologists. "It took us about four years to get that paper published. It got rejected so many times because people were just so uncomfortable with the idea of using sound to understand new physics."

Henry is eager to do more with the concept, including a sonified-periodic-table mobile app and more summers with Atlanta-area high school teachers in the lab to help convert atomic vibrations to sounds, "a fun part of the NSF grant. We're making progress. We have 15 or 20 elements so far."

BIOINFORMED DIAGNOSES

At Bristol-Myers Squibb, Ariella Sasson develops genetics-based computer analyses that help clinicians tailor treatments to patients.

By Andy Boyles

Ariella Sasson remembers well one highlight of her career. As a staff scientist at Children's Hospital of Philadelphia (CHOP), she helped develop a new clinical diagnostic test for Noonan syndrome, which causes a range of heart defects, bleeding problems, skeletal malformations and other conditions. The new assessment sidesteps the typical battery of tests for the syndrome and goes straight to the genetic mutations at fault.

To develop the test, the CHOP team used next-generation sequencing (NGS) methods on patients' DNA samples. They decoded genes previously linked to Noonan syndrome, then used computers to find activity-hindering mutations in those genes. Sasson and the bioinformatics team built a vital part of the analysis: NGS pipelines, chains of processing functions arranged so that the output of each element is the input for the next. The test quickly diagnosed the syndrome. "The head of the diagnostics lab could point to kids and say, 'You helped these kids,'" Sasson recalls. "That was pretty amazing."

Sasson is determined to work on mathematics with practical applications. With support from a Department of Energy Computational Science Graduate Fellowship (DOE CSGF), she developed the skills needed to computationally analyze NGS results. First at CHOP and now at Bristol-Myers Squibb Company, she has helped devise gene-centered tests that improve diagnosis and treatment of disorders in children and adults.

As an undergraduate at Rutgers University, Sasson earned a triple major in mathematics, biomathematics and computer science in 2001. She was discouraged to find that most applied mathematics graduate programs were small and narrowly focused. She went to work as an actuarial analyst for the Chubb Group of Insurance Companies, which had the appeal of keeping her close to family in New Jersey.

Later, she learned of Rutgers' BioMaPS Institute for Quantitative Biology. "It brought in this breadth of scientists who were affiliated with other departments into a localized environment, so you could access them all," she says.

'It's like an art project. You never know when you're finished.'

Sasson was accepted to BioMaPS and began a doctoral project in chemistry in 2004. She earned the DOE CSGF based on that work. "I have to say the CSGF was one of the best things that came my way in grad school," she says. The program relieved her of teaching responsibilities and offered the flexibility to study the subjects she chose.

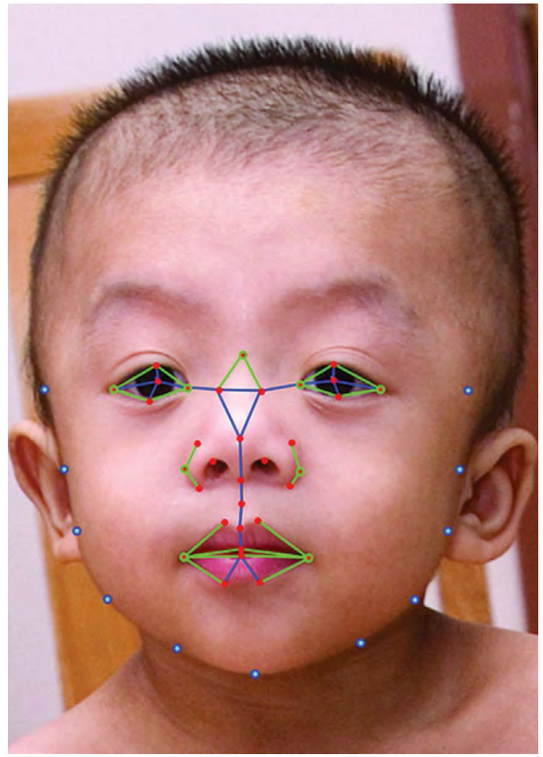
Later, Sasson attended a talk by Rutgers biologist Todd Michael about NGS. "Oh, that's really cool," she thought. Soon, she was working with Michael and with biophysicist Anirvan Sengupta as her advisor on computational analysis of NGS results. "I don't know if I would have been able to do the sequencing piece if I didn't have that funding" from the DOE CSGF, she says.

For her DOE CSGF practicum, Sasson went to Sandia National Laboratories in Albuquerque to work with W. Michael Brown on support vector machines. "It's a way to classify data," Sasson says. "You're putting a plane between different data sets." Since then, she has applied the method to genetic mutations in an attempt to identify the causes of complex diseases.

At Rutgers, her thesis centered on genome assembly, in which multiple DNA fragments from a cell or organism are matched and strung together into whole chromosomes. Assembly involves frustrations, however. "It's like an art project," Sasson says. "You never know when you're finished."

Sasson was glad to move from assembly into diagnostics at CHOP after graduating in 2010. The Noonan syndrome test proved the value of NGS-based diagnostics, and the researchers grew the project until they were scouring individual patients' genomes for several genetic disorders. Their search for the causes of congenital heart disease may soon advance diagnostics and treatment. Having found no single mutation that causes any form of the disease, they are exploring the possibility that mutations in several genes may have a cumulative effect in disrupting cardiac development.

In 2015, Sasson moved to Bristol-Myers Squibb in Pennington, New Jersey, where she works on biomarkers and stratification strategies that guide immune-system therapies. These objective measurements and procedures match patients to promising treatments to increase the probability of positive outcomes. For example, one assay measures how many mutations reside in the DNA of a patient's tumor cells. These measures of tumor mutation burden (TMB) are being evaluated as treatment guides for various cancers. Early clinical results suggest certain therapies may be effective for small-cell lung cancer patients who have high TMBs.



National Institutes of Health researchers and their collaborators plot distances and angles between different landmarks on the faces of children throughout the world to determine whether they have Noonan syndrome, a common genetic disease. *Credit: Darryl Leja, National Human Genome Research Institute.*

Since her work on Noonan syndrome, Sasson's career has hit several high points, and she looks forward to more. "I've helped people, not directly but indirectly," she says. "It's my little bit of helping to make the world a better place."

GENETIC OFF-SWITCH

With computing, Livermore's Jeff Drocco seeks ways to reverse possible genetic perils.

By Thomas R. O'Donnell

The issues Jeff Drocco tackles at Lawrence Livermore National Laboratory (LLNL) are like the woes the mythical Pandora released when she opened the fabled box.

Drocco works with the lab's Global Security Program and focuses on biodefense and biosafety – countering intentional or inadvertent releases of harmful organisms or toxins. In either case, he says, authorities face the same problem: "Once a biological agent is out there, it often takes considerable work to stop the threat."

Drocco, a Department of Energy Computational Science Graduate Fellowship (DOE CSGF) recipient from 2004-08, has proposed a way to reverse damage from a particular type of biological release with only a small amount of antidote.

His research got a boost from a 2015 fellowship with the Emerging Leaders in Biosecurity Initiative, operated by the Johns Hopkins Center for Health Security. The program brings together professionals through a series of events and webinars. At a three-day conference in Washington, Drocco learned about biosecurity developments and policies. The group later went to the United Kingdom to visit the Defence Science and Technology Laboratory and the Pirbright Institute, which studies animal diseases.

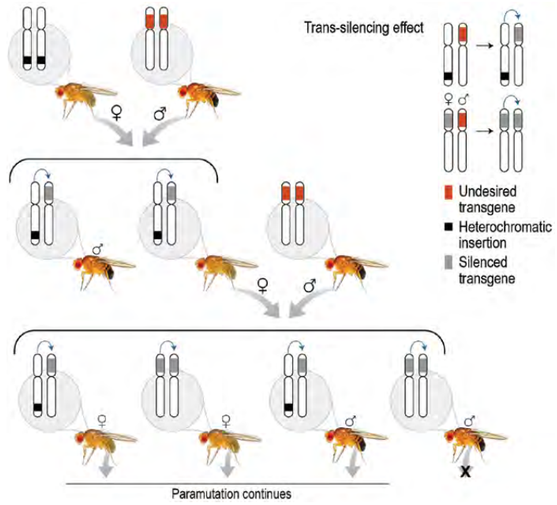
Drocco says conversations with the other biosecurity fellows helped him realize "the problems everybody else in this field are running into are as difficult as the ones (I'm) running into." He's built lasting connections with others willing to share their expertise.

One of Drocco's latest investigations involves *Drosophila*, a genus of fruit fly that is a common model organism for science and also was the subject of his doctoral thesis.

The task, supported by an LLNL laboratory-directed research and development grant, focuses on gene-drive schemes, which are designed to rapidly spread an engineered genetic change through a wild population of organisms. For example, mosquitoes could be genetically modified to resist the malaria parasite and then released into the wild. With a gene drive, they would mate with wild mosquitoes and pass the mutation to nearly all their offspring, not just the usual half.

'Any time you have these complex environmental releases, computing is always going to come into play.'

But if scientists were to produce a gene drive that succeeded in the wild, they wouldn't have a good way to stop it – if, suppose, an engineered organism were to escape from a laboratory. In an LLNL report, Drocco and Shawn Little of the University of Pennsylvania proposed a genetic off-switch and computationally calculated its effects using *Drosophila* as an example.



Trans-inactivation of a gene-drive-transmitted genetic alteration. The black region indicates the genetic insertion responsible for producing piRNAs complementary to the gene drive region, called an allele. Red and gray regions denote the gene drive allele in its active and silenced states, respectively. Inactivation spreads quickly because the silencing effect can be passed via the female egg cytoplasm even in the absence of the original genetic insertion. *Credit: Jeff Drocco.*

The technique enlists Piwi-interacting RNA (piRNA), a molecule that may help protect the genetic code in insects from undesirable changes. Drocco and Little suggest altering a gene in female flies to encode piRNA that is complementary to the gene-drive mutation. Offspring of these females and mutated males would generate piRNA that silences the gene-drive changes. The piRNAs also can pass to offspring via cytoplasm, material outside the nucleus in fruit fly eggs, helping to spread the gene drive countermeasure even more efficiently.

"For the most part, this insect is just a normal insect, though it does have that cassette of genes that are engineered in the laboratory" and spread via a gene drive. "The piRNA system provides a way of silencing it that can also spread through the population." The system is called a paramutation because it doesn't alter the target organism's genetic code. It simply blocks expression of the undesirable gene.

Drocco and Little wondered how sensitive the gene drive system was to when and how many piRNA paramutation flies are released. Their computational models found it wouldn't take an overwhelming number of piRNA-altered flies released at a particular time to effectively silence gene-driven alterations. In fact, simulations suggested the paramutation would spread just as efficiently as the targeted gene-drive mutation, Drocco says. He and Little hope to test the conclusions with experiments.

It would have been nearly impossible to even suggest such an outcome without simulation, Drocco says. "Any time you have these complex environmental releases, computing is always going to come into play because these aren't always elegant problems mathematically to figure out."

In 2017, Drocco joined the BioWatch program, created after the 9/11 attacks to thwart bioterrorism.

"I get to think about some amazingly difficult but also important problems every day," he says. Especially with BioWatch, "we're a part of some amazing systems that aren't always widely known or widely appreciated but do a lot for advancing science and protecting the country."

BIG SURPRISES COME IN NANOSCALE PACKAGES

The DOE CSGF stages the Communicate Your Science & Engineering contest to give fellows and alumni an opportunity to write about computation and computational science and engineering for a broad, non-technical audience. The author of this year's winning essay is a fourth-year fellow studying mechanical engineering at the Massachusetts Institute of Technology.

By Gerald J. Wang

At a picnic in the dog days of July, there are few joys as simple as a cup of cool water. And to the naked eye, such a drink, resting on a park bench, might seem just that uncomplicated – calm, motionless, perhaps even boring.

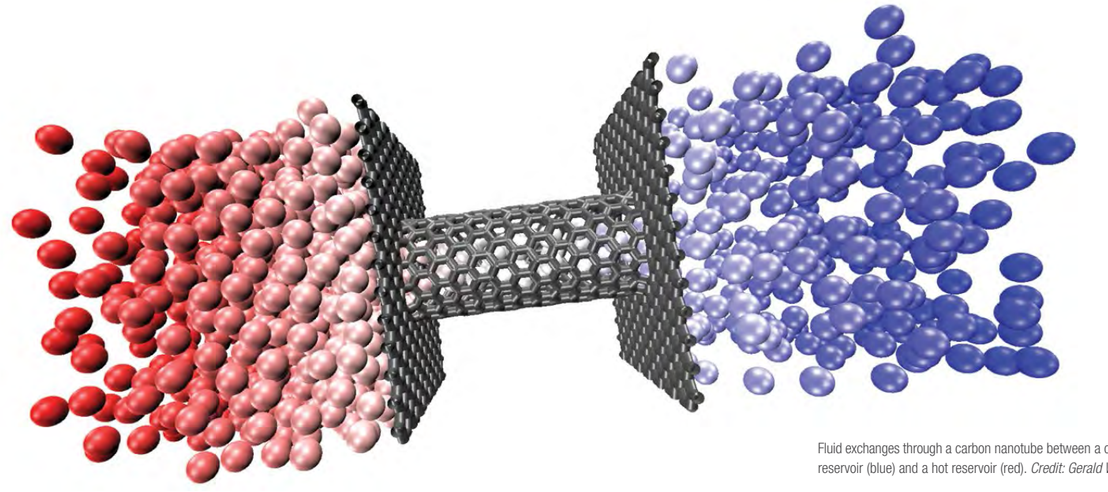
But for a scientist who thinks about everything at the atomic scale, nothing is simple about this system. If our eyes could zoom in on the water by a factor of a billion, to see the nanoscale world, we would witness a teeming swarm of molecules frantically zipping past one another like bugs at sunset. These molecules do not form lasting relationships; no two remain neighbors for any appreciable time.

Yet not all systems are so frenetic at the nanoscale. If we step indoors to place that same cup of water in the freezer, we would soon find that the molecules within ice are decidedly tame. With little desire for adventure, they stay neatly arranged in a repeating pattern, or crystal structure, like the squares on a checkered picnic blanket. At the nanoscale, the essence of a fluid is constant chaos; the hallmark of a solid is order and calm.

Now imagine we are back at the picnic table, sipping from another cup of water through a supremely slender straw – so thin, in fact, that only a few water molecules fit across its diameter. Engineers are actually making cylinders this small (called nanotubes) out of a range of materials. Simulations my colleagues and I run on powerful computers have shown that water molecules can readily flow through these extremely narrow straws – in other words, they behave like a fluid. And yet, with less opportunity to zip and swarm, these molecules will settle into a neatly repeating pattern even as they move along – a geometric compromise necessitated by close quarters. This structure, which shares many similarities with that of ice, can persist even at temperatures as high as those of a hot summer day.

In this way, molecules confined in minuscule spaces can blur the line between fluid and solid. In so doing, these materials also unlock a new world of engineering possibilities and may help address many of today's great challenges.

For example, although we might take a cup of clean water for granted, hundreds of millions of people around the world struggle to find that very thing every day. We can use our knowledge of highly confined fluids to design channels that let water molecules pass freely but are unfriendly to foreign chemical species – just like a highly ordered crystal, which resists the introduction of dissimilar atoms. Water filters made from such channels would be a boon to desalination efforts around the world.



Fluid exchanges through a carbon nanotube between a cold reservoir (blue) and a hot reservoir (red). *Credit: Gerald Wang.*

These technologies bear great promise, but designing them demands extreme precision. How can we possibly study the inner workings of such minuscule systems with such unfathomable exactness? My computational tool of choice is molecular-dynamics (MD) simulation, a technique that uses high-performance computing to tackle the following question: Given a starting configuration for a system of atoms, along with information about how these nanoscale particles interact, how can we determine their positioning at a later time?

The answer is surprisingly simple. Just as Newton's laws dictated the motion of the apocryphal apple that landed atop his head, so too can these mathematical rules be used to predict the movements of individual atoms. Using past as prologue (or initial conditions, as a computational scientist is wont to say), an MD simulation advances a nanoscale system forward in time by calculating the forces each atom exerts on every other atom. The simulation then uses these forces to determine future positions and velocities for each atom. Just as a softball player can predict the trajectory of a fly ball traveling through Earth's gravitational field, a computational scientist can predict the paths of millions of "fly atoms" traveling in a much more complicated field (one the atoms themselves generate). By analyzing these atomic tracks, we can infer many critical engineering properties, such as how quickly a fluid flows through a nanoscale straw. The ability to predict these qualities is vital as we continue to develop and refine engineering applications for highly confined fluids.

By giving us a window into the nanoscale world, MD simulation allows us to meticulously understand and control – at an atomic level – the chaos within the calm. Through painstaking nanoscale engineering, we can craft tiny solutions to solve some of the world's biggest problems. Though this line of work is no picnic, it sure seems like one.

CLASS OF 2018

Alnur Ali
Carnegie Mellon University
Machine Learning
Advisor: Zico Kolter
Practicum: Lawrence Berkeley National Laboratory
Contact: alnurali@gmail.com

Julian Kates-Harbeck
Harvard University
Physics
Advisor: Michael Desai
Practicum: Princeton Plasma Physics Laboratory
Contact: juliankh91@gmail.com

Jay Stotsky
University of Colorado
Applied Mathematics
Advisor: David Bortz
Practicum: Lawrence Berkeley National Laboratory
Contact: jay.stotsky@colorado.edu

Thomas Anderson
California Institute of Technology
Applied and Computational Mathematics
Advisor: Oscar Bruno
Practicum: Lawrence Livermore National Laboratory
Contact: tanderson@caltech.edu

Alex Kell
Massachusetts Institute of Technology
Computational Neuroscience
Advisor: Josh McDermott
Practicum: Lawrence Berkeley National Laboratory
Contact: alexkell@mit.edu

Mukarram Tahir
Massachusetts Institute of Technology
Computational Materials Science
Advisor: Alfredo Alexander-Katz
Practicum: Argonne National Laboratory
Contact: mtahir@mit.edu

Hannah De Jong
Stanford University
Genetics
Advisor: Euan Ashley
Practicum: Lawrence Livermore National Laboratory
Contact: hnd@stanford.edu

Aditi Krishnapriyan
Stanford University
Condensed Matter Physics, Materials Theory
Advisor: Evan Reed
Practicum: Los Alamos National Laboratory
Contact: aditik1@stanford.edu

Thomas Thompson
Harvard University
Geophysics
Advisor: Brendan Meade
Practicum: Oak Ridge National Laboratory
Contact: t.ben.thompson@gmail.com

Hilary Egan
University of Colorado
Astrophysics
Advisor: David Brain
Practicum: National Renewable Energy Laboratory
Contact: hilaryye@gmail.com

Ryan McKinnon
Massachusetts Institute of Technology
Physics
Advisor: Mark Vogelsberger
Practicum: Lawrence Berkeley National Laboratory
Contact: ryanmck@mit.edu

Gerald Wang
Massachusetts Institute of Technology
Mechanical Engineering
Advisor: Nicolas Hadjiconstantinou
Practicum: Argonne National Laboratory
Contact: jerry.wang@mit.edu

Kyle Felker
Princeton University
Applied and Computational Mathematics
Advisor: James Stone
Practicum: Princeton Plasma Physics Laboratory
Contact: kfelker@math.princeton.edu

Danielle Rager
Carnegie Mellon University
Neural Computation
Advisor: Brent Doiron
Practicum: Lawrence Berkeley National Laboratory
Contact: drager@andrew.cmu.edu

Kathleen Weichman
University of California, San Diego
Physics
Advisor: Alexey Arefiev
Practicum: Los Alamos National Laboratory
Contact: kweichman@ucsd.edu

Morgan Hammer
University of Illinois at Urbana-Champaign
Theoretical Chemistry
Advisor: Yang Zhang
Practicum: Los Alamos National Laboratory
Contact: mhammer@illinois.edu

Adam Riesselman
Harvard University
Bioinformatics and Integrative Genomics
Advisor: Debora Marks
Practicum: Lawrence Berkeley National Laboratory
Contact: ariesselman@g.harvard.edu

Joy Yang
Massachusetts Institute of Technology
Computational and Systems Biology
Advisor: Martin Polz
Practicum: Lawrence Berkeley National Laboratory
Contact: yangjy@mit.edu

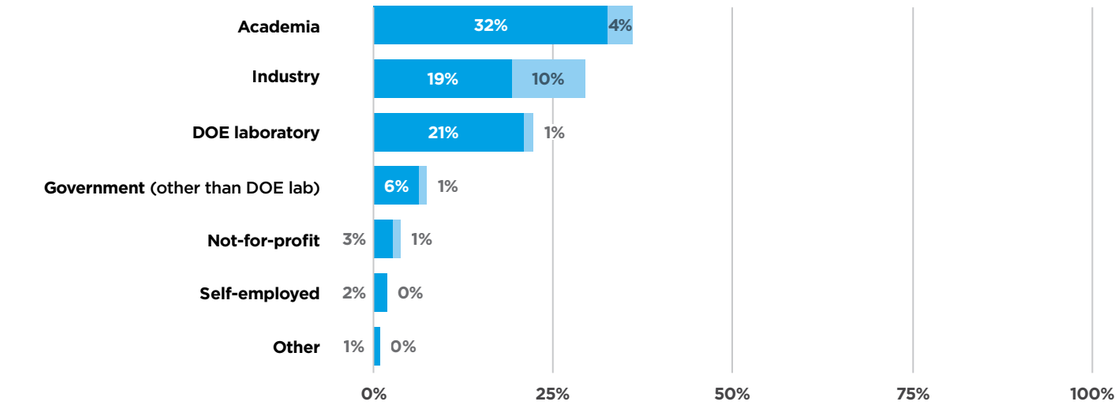
Jordan Hoffmann
Harvard University
Applied and Computational Mathematics
Advisor: Chris Rycroft
Practicum: Lawrence Berkeley National Laboratory
Contact: jhoffmann@g.harvard.edu

Adam Sealfon
Massachusetts Institute of Technology
Computer Science
Advisor: Shafi Goldwasser
Practicum: Lawrence Berkeley National Laboratory
Contact: asealfon@mit.edu

DOE CSGF AT WORK

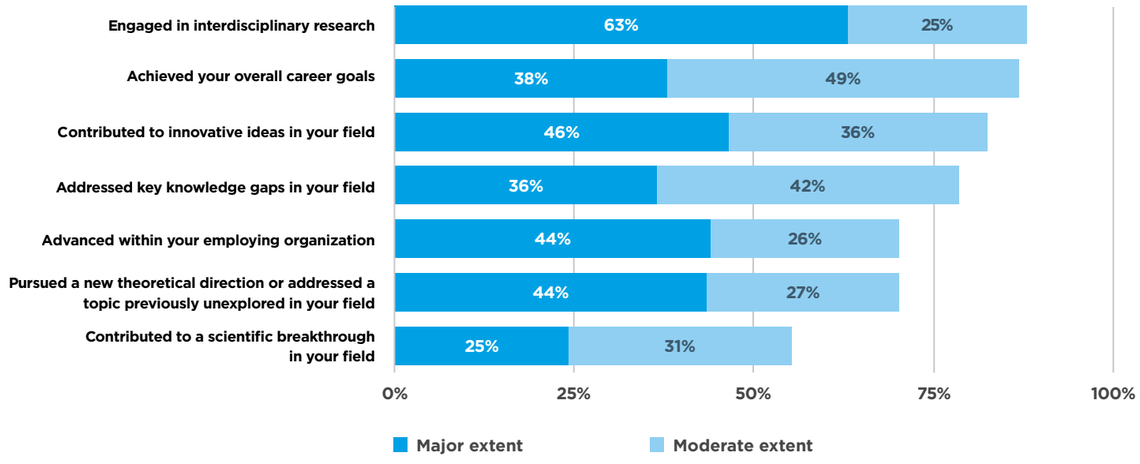
The more than 400 alumni of the Department of Energy Computational Science Graduate Fellowship (DOE CSGF) have gone on to leadership roles in government laboratories, academic institutions and private companies. They are the nucleus of a professional community dedicated to applying computing power to difficult problems. A recent survey and curriculum vitae review of nearly 300 graduates and current fellows found that the vast majority of alumni respondents (84 percent) worked in a computational science and engineering (CSE) field across a range of professional settings. Nearly all also engaged in satisfying activities such as innovative and interdisciplinary research. For more information, go to https://www.krellinst.org/csgf/about-doe-csgf/2017-longitudinal-study.

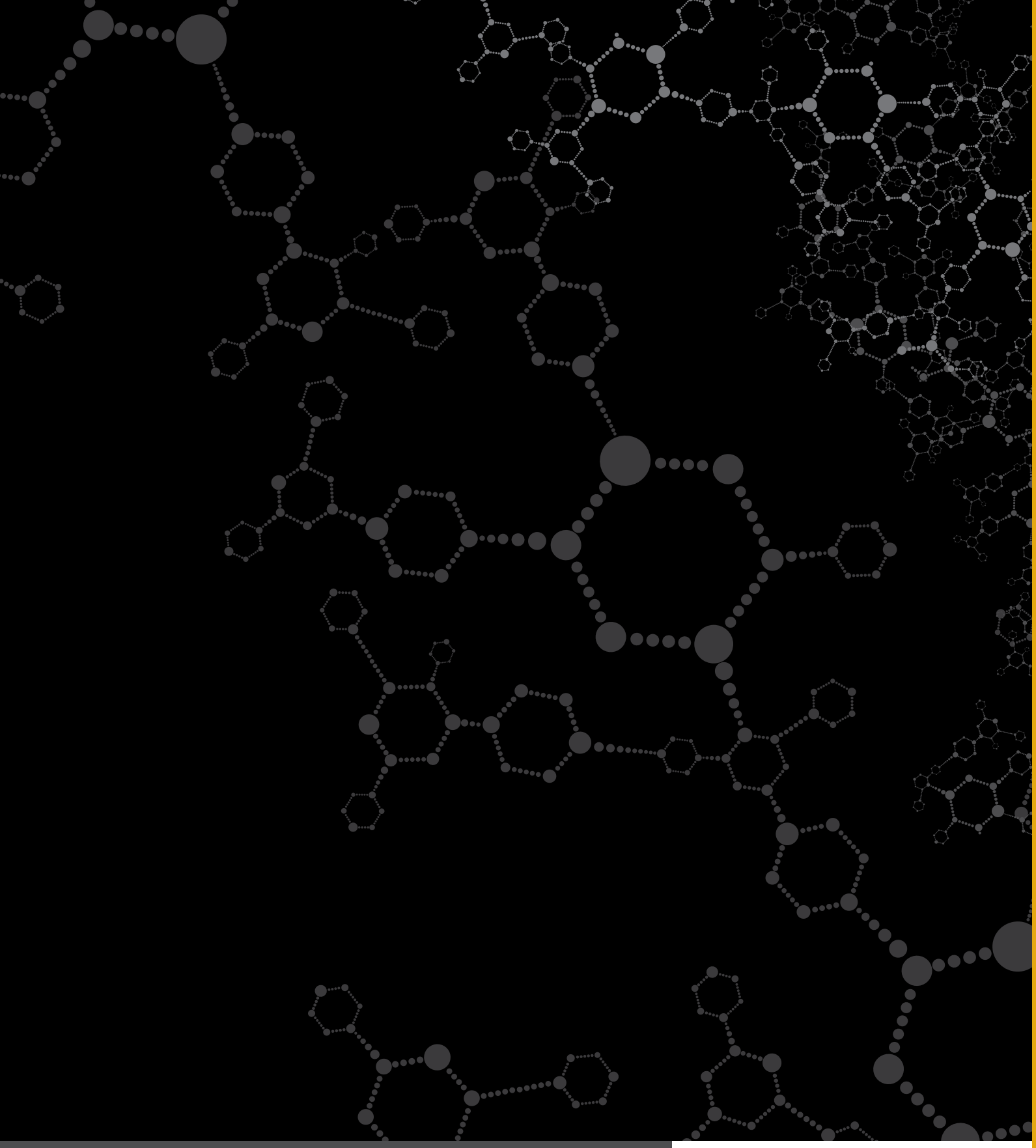
Percent of alumni employed in a CSE field, by professional setting (N=195)



NOTE: Figure is limited to alumni indicating that they were currently employed.

Percent of alumni reporting the extent to which they engaged in professional activities (N=211)





The Krell Institute
1609 Golden Aspen Drive, Suite 101
Ames, IA 50010
(515) 956-3696
www.krellinst.org/csgf

Funded by the Department of Energy Office of Science
and the National Nuclear Security Administration.