REVIEW
2019-20

Institute for Data, Intelligent Systems & Computation (I-DISC)
A Lehigh University Interdisciplinary Research Institute
The COVID-19 pandemic has highlighted the importance of data, intelligent systems, and computation. Data dashboards have become a mainstay of pandemic-related communications, from the Johns Hopkins global COVID-10 map to the New York Times US case count to Lehigh’s own coronavirus dashboard. Scientific data collected in laboratories, hospitals, and biotech firms have been disseminated with astonishing speed and openness. These data have been critical in the identification of COVID-19 hotspots and the rapid development of vaccines.

At the same time, the gaps in our data have been put into sharp relief. COVID-19 cases and deaths are likely underreported, especially in vulnerable communities, propagating further inequities in treatment. A lack of visibility into the supply chain for vaccines (and toilet paper!) has led to shortages and confusion. COVID-19 data have been misused and misrepresented, through carelessness or malicious intent, complicating the transformation of data into information. Every type of data is susceptible to misuse, but the mishandling of COVID-19 data has particularly severe consequences.

Rapid innovations in intelligent systems have assisted decision makers in the time of COVID-19. Mobile apps have facilitated contact tracing and health monitoring. Machine learning and artificial intelligence have been used to improve screening and treatment. Thanks to intelligent systems, we can schedule contactless appointments with our doctors using tele-health systems; we can order contactless delivery of food using grocery-delivery apps; we can make contactless purchases using mobile payment apps.

But “intelligent” systems, or their misuse, have also been responsible for significant failures. Stanford Medicine blamed an algorithm for the fact that many front-line workers were passed over in the first wave of vaccine distribution. The UK’s National Health System spent £22 billion on a test-and-trace system, only to scrap it due to low utilization rates and a failure to draw on international expertise.

Computational resources have been marshalled in the effort to fight the coronavirus. The 6.8 million cores in the COVID-19 High Performance Consortium have supported projects related to aerosol dispersion, population vulnerability, and drug discovery. Computational immunology is credited for the record speed at which successful vaccines have been produced. Researchers are using supercomputers to build high-resolution models of the 200 million atoms on the exterior of the SARS-CoV-2 virus.

But computational power doesn’t exist in a vacuum. It relies on data, and science, and human expertise. The potential societal benefits of this research will be tempered by our willingness to listen to the recommendations of the experts who build and analyze the computational models, and by policy makers’ ability to communicate through a fog of misinformation, conspiracy theories, and bias. Like all great human endeavors, the fight against COVID-19 is an interdisciplinary one.

Researchers at Lehigh’s Institute for Data, Intelligent Systems, and Computation (I-DISC) are tackling a wide range of interdisciplinary challenges. These scholars are engineers and computer scientists, chemists and biologists, economists and sociologists, journalists and psychologists, mathematicians and epidemiologists. They come from 22 departments, representing all five of Lehigh’s colleges.
I-DISC faculty members are forecasting COVID-19 outbreaks and tracking tumor cells. They are making algorithms more humane. They are making robots more dexterous and perceptive. They are using quantum computers for complex optimization problems. They are optimizing power grids and mining medical data. They are ensuring the resilience of critical infrastructure. They are developing computational methods for identifying the building blocks of porous materials. They are rebuilding trust in visual media and combatting human trafficking. And they are building supercomputers to power it all.

The challenges faced in the COVID-19 pandemic highlight the importance of interdisciplinary discovery, problem solving, and application. I-DISC brings together scholars from diverse disciplines and supports the process of addressing a range of real-world challenges. We hope the pages that follow will convey our excitement about these projects, and about the other events and initiatives sponsored by I-DISC. We are proud of our accomplishments and continue to look ahead to the work that remains to be done.

Larry Snyder
Industrial and Systems Engineering, I-DISC Director
Kate Arrington
Psychology, I-DISC Associate Director
Brian D. Davison
Computer Science and Engineering, I-DISC Associate Director

NEW I-DISC LEADERSHIP TEAM
I-DISC has a new leadership team, effective September 2020. Prof. Larry Snyder, Department of Industrial and Systems Engineering, was appointed Director of I-DISC, having previously served as Co-Director. Joining Larry are Prof. Kate Arrington, Department of Psychology, in the College of Arts and Sciences, and Associate Prof. Brian Davison, Department of Computer Science and Engineering, both of whom previously served on the Faculty Advisory Council. Profs. Héctor Muñoz-Avila and Katya Scheinberg, founding Co-Directors of I-DISC, have taken on new roles. Prof. Muñoz-Avila is now a Program Director for Information Integration and Informatics (III) at the National Science Foundation (NSF). Prof. Scheinberg is a faculty member in the Operations Research and Information Engineering department at Cornell University.
I-DISC BY THE NUMBERS

**I-DISC MEMBERS**
- Number of affiliated faculty members: 87

**DEPARTMENTS**
- Number of departments represented by faculty members: 22

**COLLEGES**
- Faculty members represent all 5 Lehigh University Colleges: 5

**RESEARCH GROUPS**
- Number of I-DISC research groups: 8

**PROPOSALS**
- Total number of external grant proposals whose PI is an I-DISC faculty member: 174

**I-DISC-SUPPORTED PROPOSALS**
- Number of proposals directly supported by I-DISC: 19

**I-DISC-SUPPORTED FUNDING**
- Total value of grants awarded to I-DISC research groups: $587 K

**TOTAL FUNDING**
- Total value of grants whose PI is an I-DISC faculty member: $10.8 M

I-DISC RESEARCH FOCUS GROUPS

- Artificial Intelligence & Machine Learning
- Data-Driven Inference & Control of Dynamic Systems
- FinTech & Blockchain
- Machine Learning in Materials Science
- Operations Research & Machine Learning
- Probabilistic Modeling
- Robotics
- Scalable Systems Software
Advances in sensing and communication technologies allow diverse data collection at massive scale, granting unprecedented visibility into system operations. Members of I-DISC’s Data-Driven Inference & Control of Dynamic Systems research group were awarded a $100,000 Lehigh University Accelerator Grant to develop ways to overhaul operational paradigms in order to ensure the safety, security and sustainability of these systems. This new initiative is headed by Parv Venkitasubramaniam (ECE) and includes faculty members Rick Blum (ECE), Paolo Bocchini (CEE), Brian Davison (CSE), Jeff Heflin (CSE), Héctor Muñoz-Avila (CSE) (until Sept. 2020), Ted Ralphs (ISE), and Sihong Xie (CSE). The group will conduct data-intensive data-driven research for inference and control in large-scale cyberphysical infrastructural networks. The pilot phase of their research will focus on problems in the smart electricity grid, where much of the data infrastructure is in already in place, and the methodologies developed can be evaluated. This group is currently searching for a Postdoctoral Fellow in Data Science for Engineering.

### Catastrophe Modeling

Members of I-DISC’s Probabilistic Modeling Research Group were awarded Lehigh’s first Research Futures Grant. The goal of this initiative is to leverage Lehigh’s strengths to establish an interdisciplinary Catastrophe Modeling research program at Lehigh, in parallel with its educational component. “Catastrophe Modeling” (or "CatModeling") is a rigorous probabilistic approach to the study of natural disasters and their consequences, with the purpose of estimating expected losses and risk. Recent trends are seeing CatModeling being applied beyond natural disasters, to other events like financial crises, political unrest, and pandemics. The I-DISC team is led by PI, Paolo Bocchini (CEE) and includes Javier Buceta (until summer 2020), Daniel Conus (Math), and Brian Davison (CSE). A postdoctoral Fellow is planned to start at the beginning of 2021 to work in Probabilistic Modeling of Regional Hazards, and a search continues for a postdoc to work on Catastrophe Modeling applied to Coastal Infrastructure and Communities (in collaboration with Rice University).
SSMCDAT2021 NSF HACKATHON
NSF AWARDS DMR-1938729 & DMR-1938734

The exponential increase in available computing power has made it possible to generate and analyze large amounts of data about physical materials. One of the leading challenges in solid-state materials chemistry is determining how best to make use of this abundance of materials data to accelerate the development of new materials-based technologies. PI Frank E. Curtis (ISE), in collaboration with the University of Utah, was awarded NSF funding to host the SSMCDAT2020 “hackathon” (postponed to 2021 due to COVID-19). Teams composed of both solid-state materials chemistry (“SSMC”) researchers and data (“DAT”) scientists will work together to apply advanced data science methods to address important, challenging problems in solid-state materials chemistry. It is anticipated that as a result of this workshop, the most successful teams may pursue additional funding and resources to build upon their successful collaboration. More information is available on the SSMCDAT Hackathon website.

NSF TRIPODS+X
I-DISC WORKSHOP SERIES ON MACHINE LEARNING
NSF AWARD DMS-1839353

Lehigh’s NSF TRIPODS+X award has enabled I-DISC to hold a series of interdisciplinary workshops on emerging interdisciplinary topics that have strong internal support at Lehigh and that fall under the general theme of “Machine Learning + X.” The workshops bring together researchers who otherwise may not interact to foster interdisciplinary communication and collaboration. This gives the communities early exposure to each other’s problems, techniques and terminology and results in faster development of novel convergent solutions to key societal and scientific challenges. The small scale of these workshops, single-track programs and collaborative environment create an interaction opportunity that cannot usually be achieved at professional conferences. I-DISC hosted the Data Science & Materials Conference in May 2019 and the Robot Learning Workshop in October 2019 (see pg. 7). A workshop on Machine Learning and Supply Chain Management is planned for 2021 (see pg. 14). The NSF TRIPODS+X program is a partnership between the NSF Directorates for Computer & Information Science & Engineering (CISE) and Mathematical & Physical Sciences (MPS).
MACHINE LEARNING & MATERIALS SCIENCE CONFERENCE (May 2019)

This workshop focused on the development and application of data science algorithms and tools to address problems in molecular and materials science and engineering (i.e., any problem spanning the length scale of atoms to bulk materials). The 2-day event brought together foundational and applied data science experts to present the latest machine learning methods and their applications in molecular and materials science and engineering. A poster session gave attendees a further opportunity to present and discuss their research areas in more depth. The workshop also encouraged interactions and networking through its roundtable set-up, breaks, and social events. More information is available on the workshop website.

ROBOTICS WORKSHOP (October 2019)

This 2-day event facilitated interactions among researchers from different disciplines interested in designing and implementing autonomous robots. The workshop aimed to inspire the research community toward new interdisciplinary directions in robotics, controls, and machine learning. The presentation of challenging and important problems, in a coherent fashion, to these communities can open up tremendous intellectual opportunities for research and attract young researchers and students to this timely and important research field. 16 presentations (from 12 universities in 9 different states, plus NSF and Google AI) represented emerging directions within the intersection of robotics, deep learning, reinforcement learning, control systems, and operations research. 150+ participants attended and the poster session showcased 26 posters from 12 different universities. More information and videos are available on the workshop website.
I-DISC RESEARCH GROUPS

AI & MACHINE LEARNING

Formerly known as Reinforcement Learning, led by Martin Takáč (ISE), this group of 9 faculty members from MSE, CSE, ISE, ECE, MEM and ChBE focuses on applying AI/ML methodologies in various engineering applications. One of the methodologies this group is focusing on is how to train AI/ML models efficiently in problems coming from physics and engineering applications. Some of the challenges includes interpretability and fast inference.

Activities include weekly seminars attended by faculty and PhD students with the aim of creating new working groups to tackle problems and apply for grants to support these efforts.

DATA-DRIVEN INFERENCE & CONTROL of DYNAMIC SYSTEMS

This is a new initiative and collaborative research group. The pilot phase will focus on problems in the smart electric grid where much of the data infrastructure is in place and the developed methodologies can be evaluated. See article on page 5.

FINTECH & BLOCKCHAIN

This group, headed by Hank Korth (CSE). He collaborates closely with Kathleen Weiss Hanley (Finance) and other faculty from the College of Business. Recently group members have been presenting their active research problems to each other—for example “Face recognition payment adoption” (Oliver Yao, DATA) and “Detecting fraud through financial statement tags” (Neal Snow, Accounting)—or problems they are interested in exploring—for example “Automation of regulation and of compliance for financial blockchains” (Korth and Hanley), and “Predicting the Effect of Industry Payments on Medicare Prescribing Patterns - A Social Media Perspective” (David Zhang, DATA).

The goal is to focus on creating new working groups to apply for Lehigh grants. The group has also collaborated with a student research team in conjunction with the Scalable Systems Software (SSS) lab.

MACHINE LEARNING in MATERIALS SCIENCE

At Lehigh, data science and machine learning tools are being used for modeling of complex molecular and material systems of physical, chemical or biological origin to discover the underlying synthesis-structure-property relationships and leverage them in inverse design of new materials and molecules. Headed by Srinivas Rangarajan (ChBE), the group's other members include faculty from MSE, ISE and MEM.
ML in Materials Science Group members Ed Webb (MEM) and Rangarajan, along with Ganesh Balasubramanian (MEM) and Lisa Fredin (Chemistry), recently won a $400,000 NSF campus cyber-infrastructure grant to purchase a new computing cluster for Lehigh to carry out interdisciplinary collaborative research at the intersection of machine learning, first principles and ab initio computation, and materials science. In addition, Rangarajan, Jeetain Mittal (ChBE) and Mark Snyder (ChBE and I-FMD) won a three-year NSF grant to design and synthesize novel multi-porous crystalline porous organic frameworks. Rangarajan and Mittal were also the organizers of the ML & Material Science workshop.

PROBABILISTIC MODELING

This group is headed by Paolo Bocchini (CEE) and consists of a group of faculty from ChBE, CSE, Mathematics, ISE and ME. The group develops innovative approaches in probabilistic analysis and in transferring methodologies from one field to others, including probabilistic regional hazard modeling, non-homogenous virus spreading, big-data analysis and infrastructure systems resilience.

All the activities of the probabilistic modeling group have been folded into the Catastrophe Modeling initiative during 2020. As part of this, they submitted two major collaborative proposals (establishment of research and a Master's Degree program in Catastrophe Modeling and research on resilience of military installations).

PROBABILISTIC MODELING
Other activities have included organizing a Professional Certification for Catastrophe Modelers (in collaboration with the Bermuda Institute of Ocean Sciences); completing the PRAISys project on Resilience of Infrastructure Systems, which involved 55 scholars, nearly half of whom are from groups that are underrepresented in STEM; conducting a project on multi-hazard quantization, funded by a Lehigh University Accelerator grant; and developing and offering a new summer course on Resilience of Systems.

ROBOTICS

This team consists of 5 faculty members from MEM and CSE, led by Nader Motee (MEM). They have been focusing on developing conceptual, computational, and mathematical methodologies for creating robots that sense and react quickly and robustly with the goal to achieve long-term robot autonomy by integrating sensing, learning, reasoning, task planning, and control. This group organized the 2nd TRIPODS+X workshop on Robot Learning.

SCALABLE SYSTEMS SOFTWARE

Led by Roberto Palmieri (CSE), this group also forms part of Lehigh's CSE Scalable Systems and Software Research Group and a number of smaller working groups within Lehigh that regularly meet. Their collaborative efforts are focused on designing and building parallel, concurrent and distributed systems. Examples include the computer science aspects of Blockchain (Korth and Palmieri); Computer Science systems research and robotics (Palmieri and David Saldaña, CSE); and a Computer Science and Business (CSB) capstone project, in which a team of four undergraduate students are working with Ethereum Classic Labs on Blockchain, under the supervision of Korth.

In addition there are working groups with external collaboration: the working group on integrating theoretical and systems aspects of distributed computing led by Palmieri and Dr. Tseng (Boston College); and the working group on combining distributed computing protocols and storage led Palmieri, Dr. Tseng, and Dr. Konwar (MIT).

Group member Michael Spear (CSE) served as the Program Chair for the 32nd ACM Symposium on Parallelism in Algorithms and Architectures (ACM SPAA 2020), and Palmieri served as the Local Arrangement Chair. Korth chaired a roundtable on Blockchain at the 2020 International Conference on Very Large Data Bases (VLDB). Finally, Xiaochen Guo (ECE) received a Department of Energy (DoE) sub-contract to perform research on "Innovative Architectures for High Performance Data Analytics".

This group was recently featured in Lehigh's Resolve Magazine in the article "Getting into Synch" (2020).
I-DISC seminars give faculty members and students opportunities to bridge Machine Learning and Data Science across disciplines. For example, Chinedu Ekuma (above left) presents at an AIML seminar; and founding I-DISC Co-Director Héctor Muñoz-Avila (above right) presents to Cognitive Science faculty & students about deep learning.

I-DISC SPONSORED EVENTS

I-DISC co-sponsored the QUANT/ Financial Engineering Symposium hosted by the Department of Financial Services and the College of Business in 2019 and 2020. These conferences aim to promote the exchange of ideas in data management, data science, Blockchain, deep learning and other areas. A number of I-DISC members have presented and participated in panel discussions.

OUTREACH

I-DISC member Martin Takáč (ISE) has developed a partnership with Saucon Valley School District and uses "Ozobots" to help engage young students in a fun and creative way to learn coding.
I-DISC continues to develop and provide services to help support Lehigh University faculty to pursue funding opportunities and forge new research collaborations across campus.

- I-DISC faculty support
- Grant-writing support
- Paper-editing services
- Event planning and support
- K-12 outreach support

More information is posted on the I-DISC website in the "I-DISC Members" tab. (Lehigh credentials required).

I-DISC CAMPUS CONNECTIONS

The inaugural Data for Impact Summer Institute, headed by the Office of Creative Inquiry and planned in collaboration with the Martindale Center and I-DISC, was an eight-week program that ran from June 15th-August 7th, 2020. 60+ students and 19 faculty project mentors collaborated on 15 projects.

Data for Impact Summer Institute

I-DISC Faculty Members mentored and supported students in the following Data for Impact projects:
- Josh Agar (MSE), Real-time Machine Learning in Experimental Materials Science
- Sihong Xie (CSE), Spam Spotting: Using AI Tools to Educate and Improve Online Decision-Making
- Lifang He (CSE), A Dose of AI for Healthcare Data Analytics
- Patrick Zoro (Finance), Falling Knives
- Srinivas Rangarajan (ChBio), Data to Control: Towards Data-Driven Model Predictive Control for Chemical Process Automation
- Ed Webb (MEM), From Molecules to Medicine: Overcoming the Time Scale Challenge
I-DISC NEW INITIATIVES

I-DISC FELLOWS PROGRAM

I-DISC announced this new initiative in December 2020 with the aim of helping Lehigh University faculty across all disciplines with the data science, machine learning, artificial intelligence, or computing part of their research.

I-DISC Fellows are advanced doctoral students with expertise in data science, computation, optimization and/or machine learning. The new I-DISC Fellows program provides a part-time research assistant to help Lehigh faculty with, for example:

- Data science / machine learning / artificial intelligence ideas, debugging and design assistance.
- Short-term software development or data science projects.
- Consulting or preliminary work for a funding proposal.
- Assessment or estimate of level of effort for a future funding proposal.

8 projects have been identified (to date) for Spring 2021.

I-DISC UNDERGRADUATE RESEARCH GRANTS

These grants encourage existing I-DISC faculty members who are part of a collaborative research group to include undergraduates in their research by providing grants between $1,000-$2,000 to pay for student stipends.

I-DISC SPONSORSHIP & SUPPORT FOR OUTREACH EVENTS

I-DISC offers $500-$1,000 awards to support I-DISC faculty for important causes, for example increasing diversity in STEM and to increase awareness of the Institute and its role within the University.

More information on these new initiatives can be found on the I-DISC website at: idisc.lehigh.edu/opportunities
UPCOMING NSF TRIPODS+X WORKSHOPS

2021
MACHINE LEARNING & SUPPLY CHAIN MANAGEMENT WORKSHOP

The vast majority of applications of ML for supply chain, both in academia and in practice, fall under the category of either descriptive or predictive analytics, which are important and relatively well-developed uses of ML in supply chains. In contrast, the focus of this workshop is the use of ML for prescriptive analytics in supply chains, i.e., for supply chain optimization. It will bring together researchers from both the ML and supply chain communities in order to foster a vibrant exchange of ideas and to stimulate new collaborations. The workshop will consist of invited speakers, a poster session for students, a panel discussion, and a one-day training event for graduate students preceding the workshop to learn the basics of supply chain management and relevant ML approaches. More information will be posted on the workshop website page.

2022

The anticipated fourth and final workshop in the TRIPODS+X series will be held in Spring 2022 or Fall 2022. The topic and details of this workshop will be announced and posted on the I-DISC website in the near future. Please check the I-DISC events page regularly for more information.

OTHER UPCOMING EVENTS

SUMMER 2022
INTERNATIONAL CONFERENCE ON CONTINUOUS OPTIMIZATION (ICCOPT 2022)

I-DISC members Frank E. Curtis and Daniel Robinson will co-organize the upcoming International Conference on Continuous Optimization (ICCOPT) to be held at Lehigh University in Summer 2022. I-DISC will be supporting this event in conjunction with RCEAS and the Industrial and Systems Engineering Department. More information on this event will be posted in the coming months.
TRANSITIONS

The I-DISC Faculty Advisory Council (FAC) includes professors from Lehigh’s colleges of engineering, business and arts & sciences who help guide the Institute’s research and collaborative initiatives.

I-DISC WELCOMES NEW FACULTY ADVISORY COUNCIL MEMBERS*

- Paolo Bocchini, Civil & Environmental Engineering, RCEAS
- Nancy Carlisle, Psychology, College of Arts & Sciences
- Frank E. Curtis, Industrial & Systems Engineering, RCEAS
- Chinedu Ekuma, Physics, College of Arts & Sciences
- Kathleen Weiss Hanley, Finance, College of Business
- Alberto J. Lamadrid, Economics, College of Business
- Roberto Palmieri, Computer Science & Engineering, RCEAS
- Parv VenKITAsubramANiAM, Electrical & Computer Engineering, RCEAS
- Yahong Rosa Zheng, Electrical & Computer Engineering, RCEAS

*New members in bold.

I-DISC wishes to thank Jeetain Mittal, Oliver Yao and Yue Yu for their contributions to the Faculty Advisory Council during their tenure in 2019-20.
I-DISC EXTERNAL ADVISORY COUNCIL

The I-DISC External Advisory Council (EAC) includes thought leaders from industry and national labs who provide input and guidance to help shape I-DISC activities and research directions.

- David Aha, The U.S. Naval Research Laboratory
- Jon Bentley, Bell Labs (Retired)
- Joseph A. Furlong, MIAC
- Lin He, Intel
- Sayten Kale, Google Research
- Michael Liebman, Bloomberg LP
- John Matranga, OSIsoft
- Jayant Kalagnanam, Thomas J. Watson Research Center, IBM
- David Stracuzzi, Sandia National Laboratory
- Mike Watson, Coupa Software
- Stefan M. Wild, Argonne National Laboratory
- Yan Xu, SAS

I-DISC STAFF

- Sarah Wing, Business Manager

STAY CONNECTED

idisc.lehigh.edu  @LehighDISC
idisc@lehigh.edu  /company/lehighdisc