NCSA Faculty Fellows 2019 Idea Acceleration Workshop

Amanda Lombardo,
Assistant Director for Research, NCSA
alombar@illinois.edu

Farzaneh Masoud,
Strategic Research Development Administrator, NCSA
fmasoud2@Illinois.edu

National Center for Supercomputing Applications
University of Illinois at Urbana–Champaign

NCSA Fellow Program

- An opportunity for faculty and researchers at the UIUC to catalyze and develop long-term research collaborations between Illinois departments, research units, and the NCSA.
- Competitive program provides seed funding for demonstration, start-up projects, workshops, and/or other activities with the potential to lead to longer-term collaborations around research, development & education.

- **Feb. 1, 2019: NCSA Fellowships Ideas Accelerator Workshop.**
  - **10 am-12 pm, NCSA Lobby. Speed match-making event between potential applicants and NCSA staff and researchers.**
  - **Now accepting proposals through NCSA Faculty Fellows website**
- **Feb. 22, 2019:** Deadline to submit to the NCSA Fellowships Program
- **April 2019:** Target date for decisions

Faculty Slides

Faculty Fellows Idea Acceleration Workshop
February 1, 2019

Combining the atmospheric and economic transmission of air pollution across countries: new evidence based on a wider set of international trade linkages and pollutants

Sandy Dall'erba

Associate Professor, Department of Agricultural and Consumer Economics

dallerba@illinois.edu

Nicole Riemer

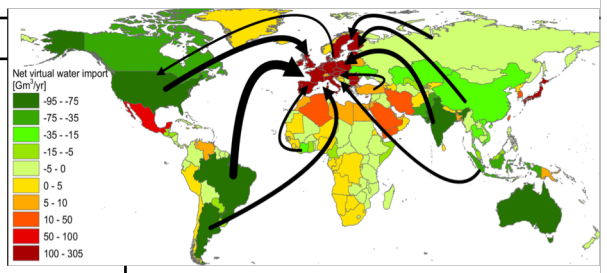
Associate Professor, Department of Atmospheric Sciences

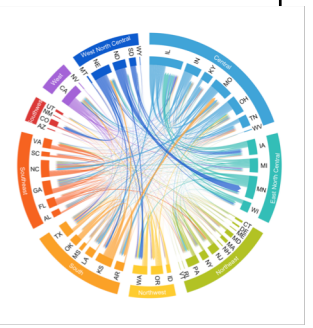
nriemer@illinois.edu

Yilan Xu,
Assistant Professor,
Department of Agricultural and Consumer Economics
yilanxu@illinois.edu

- International transmission of air pollution through trade (input-output)
- International transmission of air pollution through the atmosphere (GEOS-Chem)
- => research question 1: compensate or cancel each other ?
- => research question 2: what pollution abatement policies ?

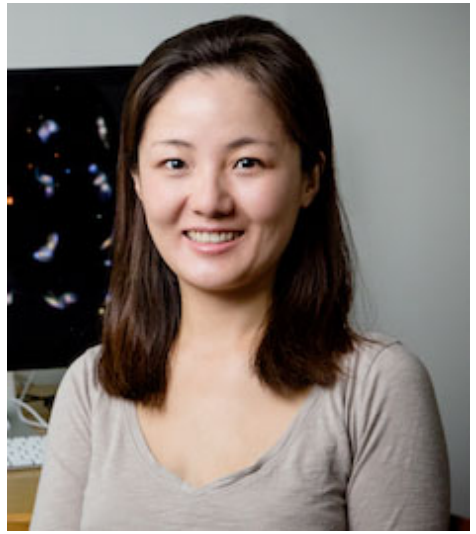
Why NCSA? => space-time-sector-altitude visualization of the flows. Our tools are static (1 D), not dynamic (4D).
NCSA DATA ANALYSIS AND VISUALIZATION GROUP or
INNOVATIVE SOFTWARE AND DATA ANALYSIS ?





- UIUC ISEE seed grant (Jan.-June 2019)
- Future: NASA (Earth Science Division Applied Sciences Program); NSF (Geography and Spatial Sciences Program; Dynamics of Coupled Natural and Human Systems)
- **NCSA seed grant to focus on visualization part**

Weighing Black Hole with Machine and Deep Learning



NCSA Thematic Area:
Astronomy & Astrophysics

Xin Liu
Assistant Professor of Astronomy
xinliuxl@illinois.edu
https://publish.illinois.edu/liu-group/

Problem

- Traditional black hole weighing methods require expensive spectroscopic observations: ~1000,000 mass measurements from 20+ yr of state-of-the-art community efforts
- The Large Synoptic Survey Telescope (LSST) will discover >1000,000,000 new black holes for which mass measurements would take 20,000+ yr with traditional methods

An artist's rendition of a supermassive black hole, such as the one at the center of the Milky Way, with an accretion disk of gas emitting light time series. (Credit: NASA/JPL-Caltech)

Key Idea

- Black hole accretion disk emits light that varies with time. Machine/Deep Learning can help decode physical information in the light time series to weigh black holes much more efficiently than traditional methods.

Project Needs

- Expertise in Machine/Deep Learning and GPU
- Supercomputing resources & data analysis infrastructure
- Success will transform the field of black hole research for fundamental science (e.g., understand the origin and evolution of supermassive black holes, enable the usage of quasars as standard candles for cosmology)

Robust projection of future urban environment

Project goal: **provide robust (multi-model) projection of global urban environment (microclimates + air quality) under climate change**

Methods:

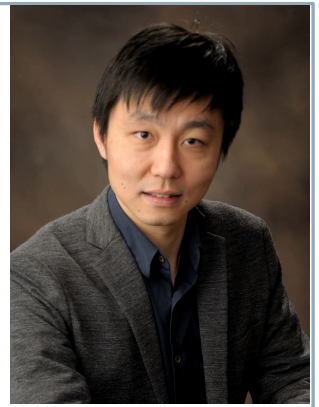
- process-based simulations + machine learning-based emulator

Resources needed:

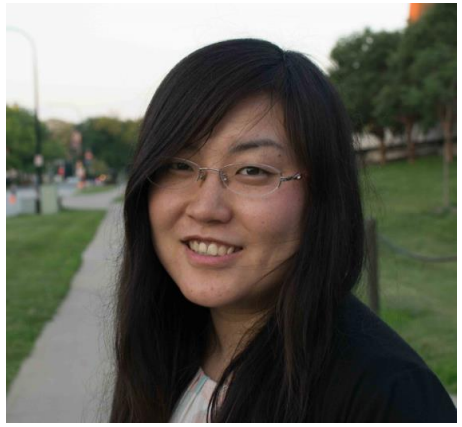
- Blue Waters or XSEDE
- Innovative Software and Data Analysis
- Data Analysis and Visualization

Contacts:

Lei Zhao
Assistant Professor
Civil and Environmental Engineering
3211 Newmark Civil Engineering Bldg
205 N. Mathews
Urbana Illinois 61801
(217) 300-9546
leizhao@illinois.edu



Source: NASA



Real-time computational framework for modeling continental-scale satellite crop phenological progress

Chunyuan Diao

Assistant Professor of Geography and GIScience

chunyuan@illinois.edu

https://bit.ly/2GcKI7Z

Project Goals:

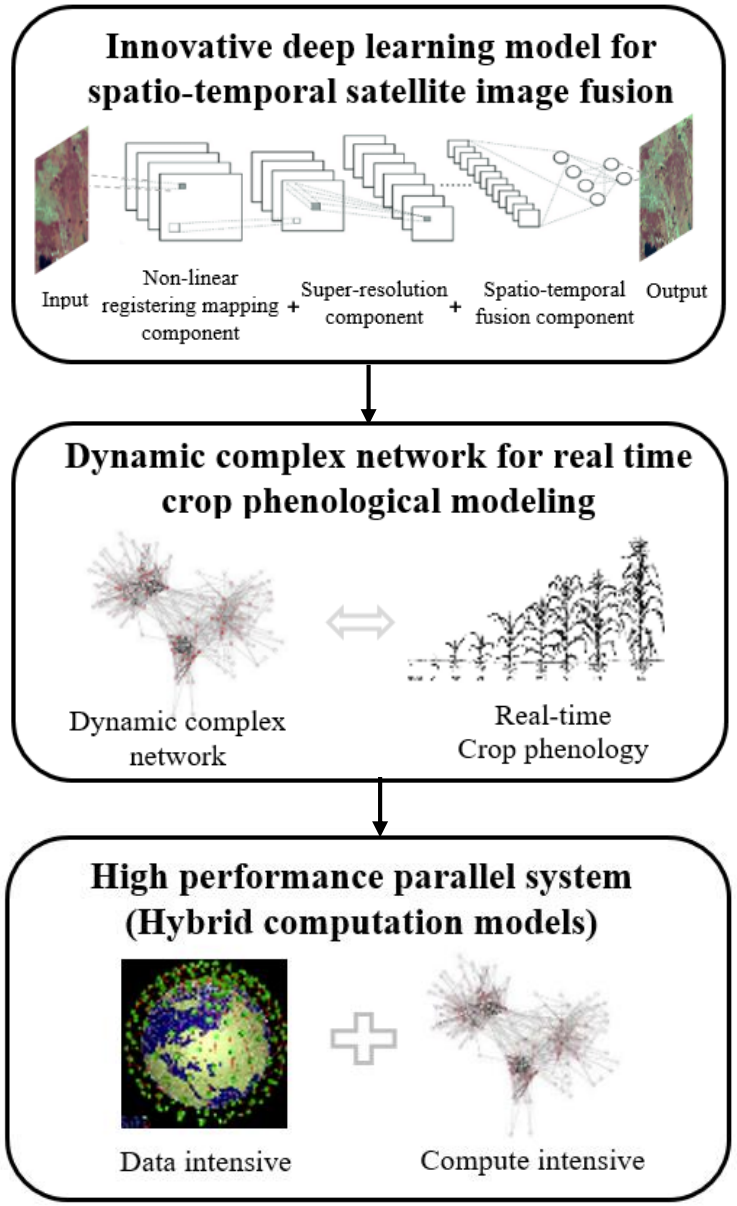
- Develop an innovative deep learning model to fuse massive satellite imagery at various spatial and temporal scales
- Develop a dynamic complex network model using the fused imagery to predict real-time crop phenological progress at continental scale
- Devise a high-performance parallel system to prototype the data- and compute-intensive satellite crop phenological modeling system

Project Significance:

- Continental-scale crop phenological modeling is crucial for assessing the influence of various farm management strategies and extreme weather conditions on crop yields
- Fusing the gigantic satellite imagery at various spatial and temporal scales is a compelling need, now more than ever, for dynamic earth system monitoring

Project Needs:

- Expertise in deep learning and parallel optimization in HPC environment
- Large-scale parallelism to handle massive satellite imagery



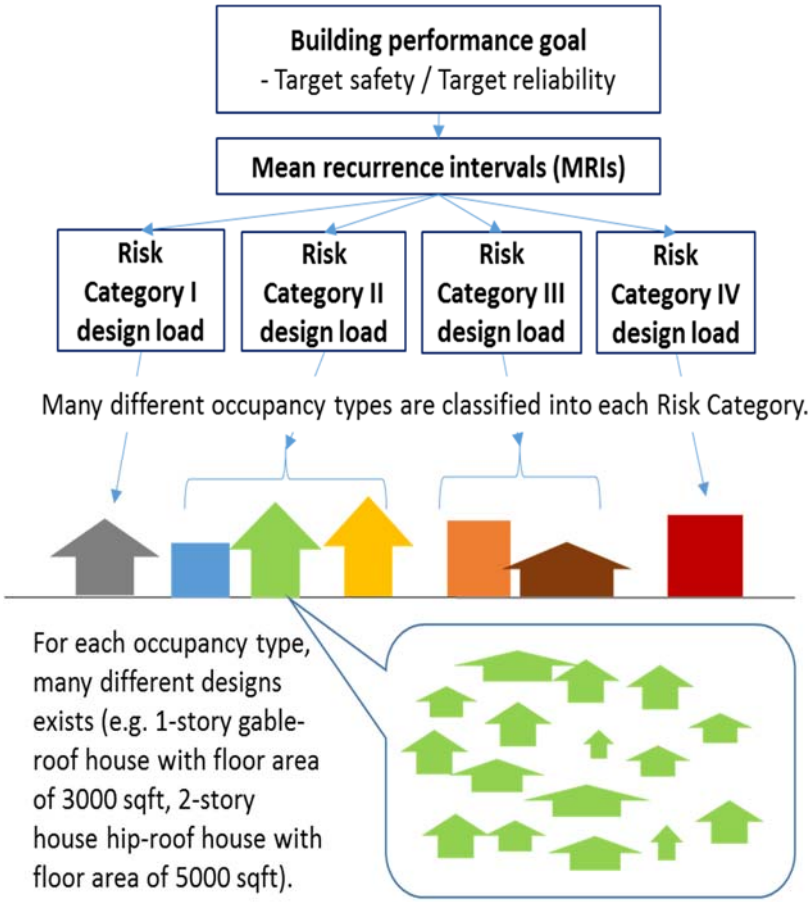
Computational Platform Development for Holistic Risk Management Framework of Built-Environment

Eun Jeong Cha

Assistant Professor of Civil and Environmental Engineering

ejcha@Illinois.edu /217-244-8343

http://publish.illinois.edu/rfouratcee/



PROJECT SYNOPSIS

- Safety targets of structural design affect overall built-environment.
- Building design safety target determination process requires structural response analysis of buildings with different risk categories, occupancies, structural types, configurations, etc, for many alternative safety targets and for many risk scenarios (MCS).
- Science-based and also practical computational tools are needed to integrate holistic risk management concept into the determination process of building design safety targets.
  – Surrogate models for structural response analysis.
- To develop surrogate models for structural response analysis, extensive amount of structural response data needs to be generated from nonlinear structural analysis.

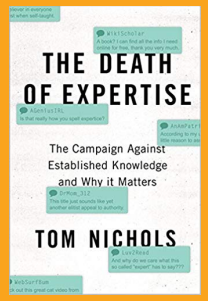
PROJECT NEEDS

- Professional support in parallel computing.

The War on Professional Expertise: The Global Spread of On-Line Myths about Medicine and Health

- **Did you know there's a cure for cancer in the basement of the White House?**
- **That malaria vaccines cause infertility in women?**
- **That HIV/AIDS is a conspiracy to kill Black people?**
- **Then you aren't keeping up! ☺**
- **This project proposes to use data mining techniques to examine the network-based spread of myths about medicine and health.**
- **(Potential collaborators: Loretta Auvil and Scott Althaus, Cline Center – funding targets: Pew; Robert Wood Johnson; NIH; SCISIP/NSF)**

"90 percent of the Information on the Web is crap…"



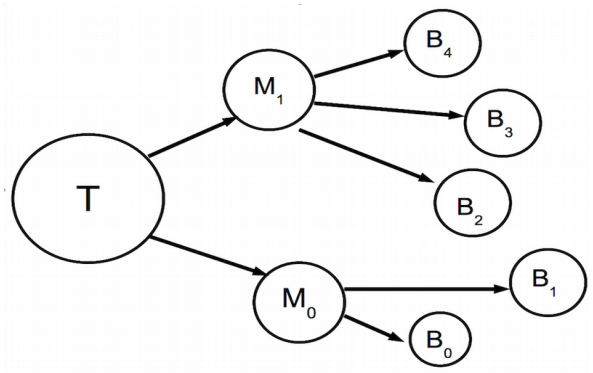
Kevin T. Leicht
Professor of Sociology
Department of Sociology
Lincoln Hall #3084
(217) 330-1399
kleicht@Illinois.edu

Composition as an Evolving Entity

Sever Tipei,
Computer Music Project,
School of Music
s-tipei@illinois.edu

Existing Software for Composition and Sound Synthesis includes:

- Randomness at multiple levels leading to multiple unique variants of a piece - an equivalence class of compositions
- It has a tree structure (Directed Graph), parent-children relationships



Goal: an Evolving Entity, a continuously changing piece

- Successive, uninterrupted computations guided through measuring: Entropy/Information, Redundancy, Complexity
- Start with few vertices → grow the structure → decay a living organism (Artificial-Life)

John Cage: "...imitate Nature in its mode of operation..."

Areas of possible collaboration

- Complex Dynamic Systems
- Information Theory/Graph Theory
- Kolmogorov Complexity
- A-Life
- Biology

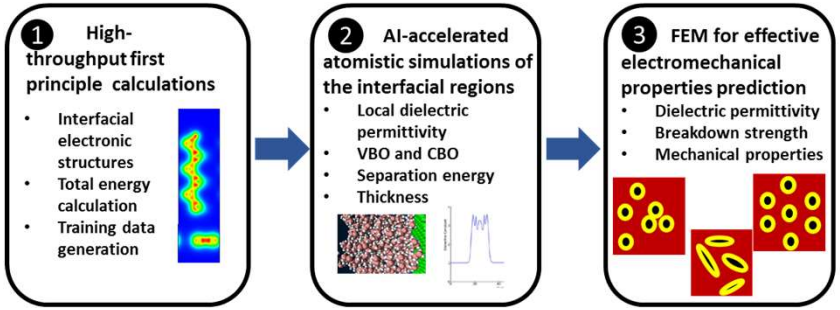
Future applications

- Sound installations
- Sonification

Pushing the Limit of Advanced Nanodielectrics through AI-accelerated Multiscale Simulation

Yumeng Li, yumengl@Illinois.edu
Assistant Professor of Industrial and Enterprise Systems Engineering

Problem: Challenges in design of Advanced Dielectric Nanocomposites (a) lacks of understanding in critical cross-scale interface effect (b) first principles computations with high accuracy but low computational efficiency and limited scalability (c) an absent unified multiscale model for systematic investigation and rational design of nanodielectrics



Needs: (a) high throughput first principle calculations (b) large-scale MD simulation parallelism (b) supercomputing infrastructure

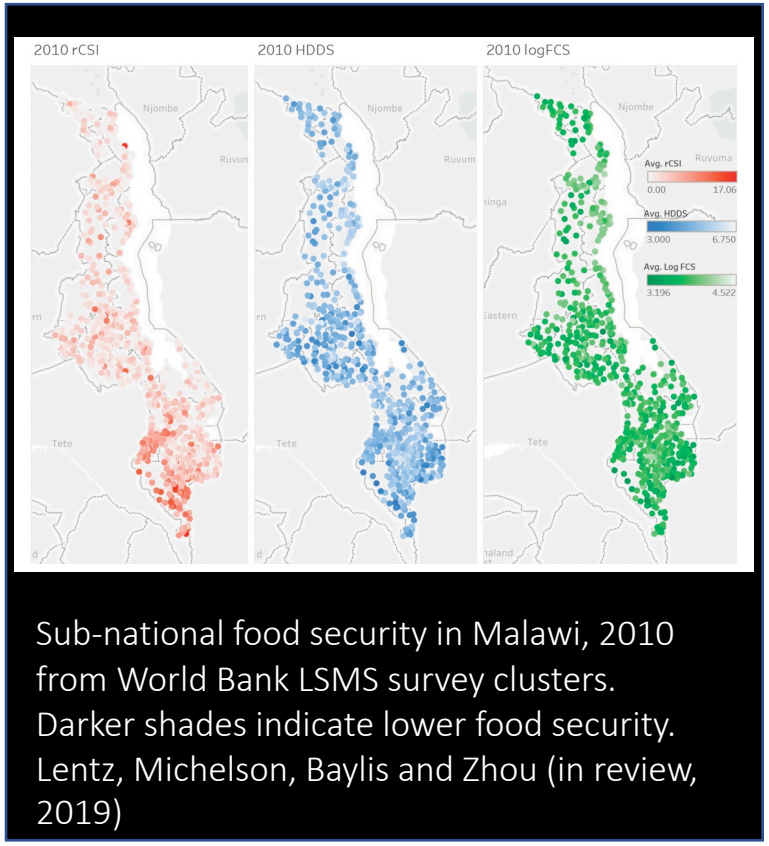
Promises: Success will accelerate the discovery of advanced dielectric nanocomposites to meet the demands in next generation electronics and electric power industries.

Approach: Establish AI-accelerated physics-based multiscale simulation framework for (a) adopting machine learning techniques for developing high fidelity interatomic potentials (b) establishing the relationship between the electronic/atomistic structures and the local interfacial dielectric, mechanical and thermal properties in the interfacial regions using atomistic simulations (c) accurate predicting effective material properties of dielectric nanocomposites, such as dielectric constant, breakdown strength and thermal conductivity with proper consideration of the interface effects

Improving famine early warning: using spatially and temporally rich data to predict food crises

Hope Michelson hopecm@Illinois.edu

Assistant Professor of Agricultural and Consumer Economics



Sub-national food security in Malawi, 2010 from World Bank LSMS survey clusters. Darker shades indicate lower food security. Lentz, Michelson, Baylis and Zhou (in review, 2019)

Problem

Need to improve prediction of food security crises

- Current food security early warning systems are ad hoc and insufficiently spatially and temporally granular
- We lack data in the places and at the times that crises are most acute

Idea

Use web crawlers and other methods to acquire new data.

Improve prediction of sub-national, real-time food security status

Need

Expertise in python, machine learning, real time prediction and visualization

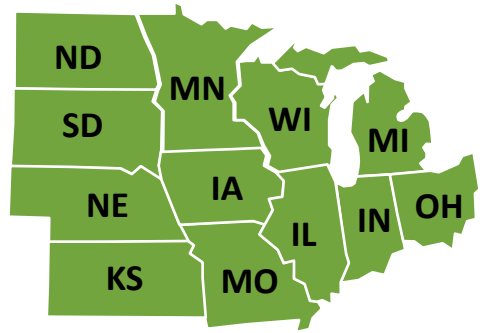
NCSA Major Programs

MIDWEST BIG DATA INNOVATION HUB

Bill Gropp, PI
Director, NCSA

Melissa Cragin
Executive Director

Co-PIs:
- Sarah Nusser, VPR @ ISU
- John Mihelich, VPR @ UND
- Brian Athey, MIDAS co-Chair @ Umich
- Inna Kouper, D2I @ IN Univ.



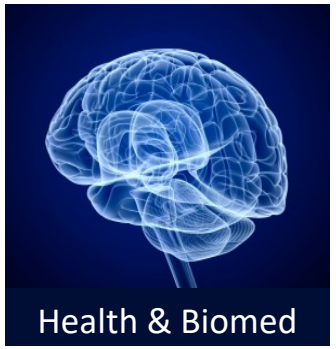
Mission

- Grow the national data ecosystem
- Develop effective partnerships across academic-industry-government-NGOs
- Build regional capacity in Data Science and Big Data

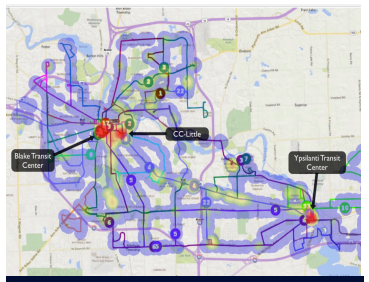
Digital Agriculture

Water Quality

Smart & Resilient Communities



Health & Biomed



Transportation

Materials & Mfg

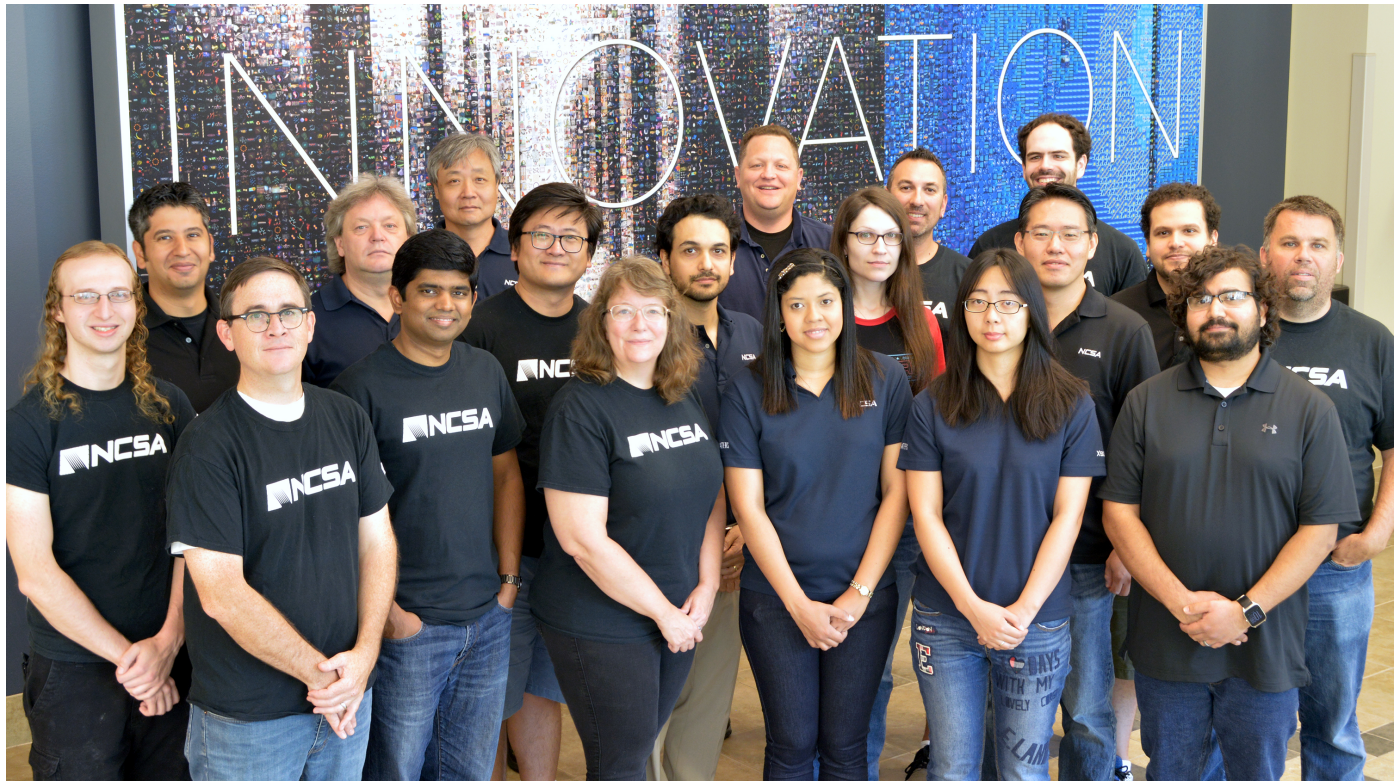
Cross-cutting areas

- Data science education and training
- Data Infrastructure: tools, CI, methods & services
- Data policy: access, use, governance

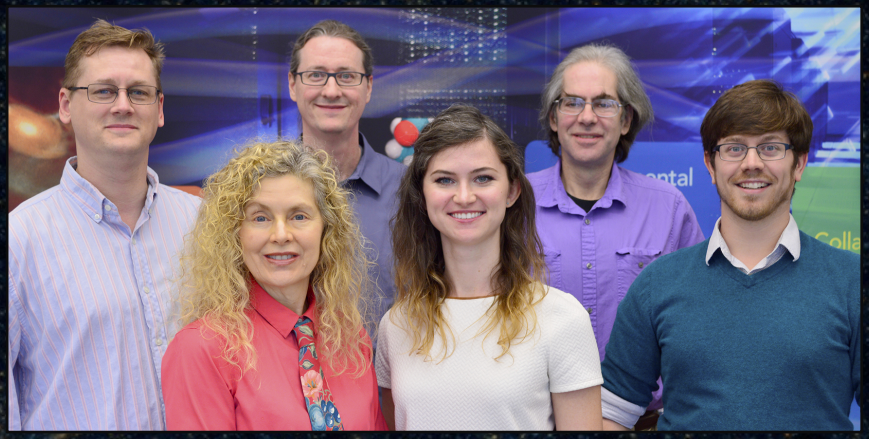
NCSA

Innovative Software & Data Analysis

- Research & Development
- Reusable software tools & frameworks for data analysis
- Bridging and amplifying efforts across different projects
- New custom software tools & frameworks



Advanced Visualization Lab



Jeff Carpenter, Bob Patterson, Stuart Levy,
Donna Cox, Kalina Borkiewicz, AJ Christensen
avl@ncsa.illinois.edu

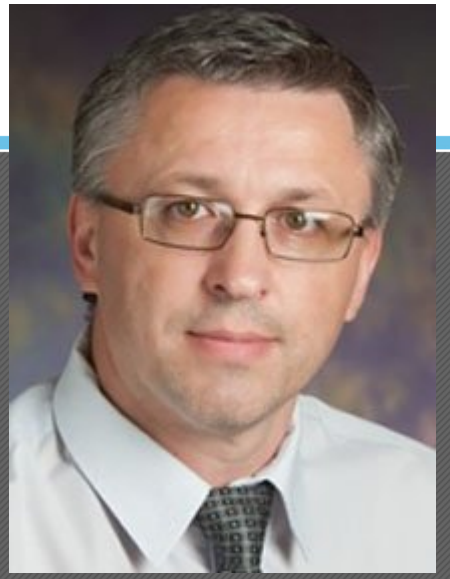
Data/Scientific Visualization

- Visualizations of computational and observational data
  - Astrophysics, geoscience, molecular biology, sustainable agriculture, geospatial systems, etc.
- Interactive visualization
- Data wrangling
- Multimedia and video production

Arts and Humanities

- Virtual Reality
- Visualization for live performances
- Mobile apps

Innovative Systems Laboratory (ISL)



kindrtnk@Illinois.edu

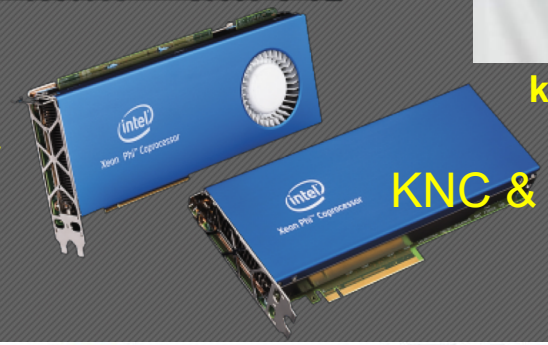
- NCSA's place for special projects
  - Investing into future projects that will result in major developments (impact, funding, …) 5+ years from now
  - Long-term support for 1-3 such projects
- NCSA's research and development vehicle
  - Primarily serve NCSA's mid-term needs in core technology areas
    - support for evaluating new technologies
    - support for developing new technologies
  - R&D agenda setup by the NCSA technical leadership with input from faculty and staff
    - 2-5 projects/technologies
    - Looking forward 1-5 years
    - Focus on new systems or impact on existing systems
- NCSA R&D business development
  - Collaboration with faculty and industry to establish new programs
  - Support for preparing proposals

Cluster of
8 x86 servers
2x P/V100 GPUs

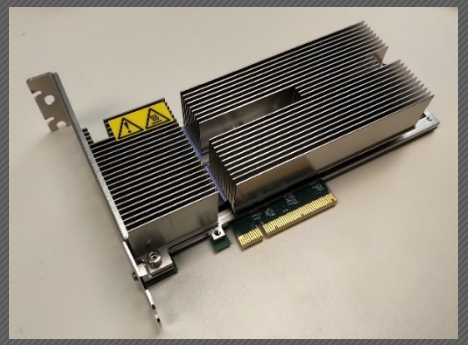
Cluster of 16 IBM
AC922 4x V100 GPUs

OpenStack cloud

CEPH storage cluster



KNC & KNL



Nallatech 250S+
CAPI 2.0 FPGA board

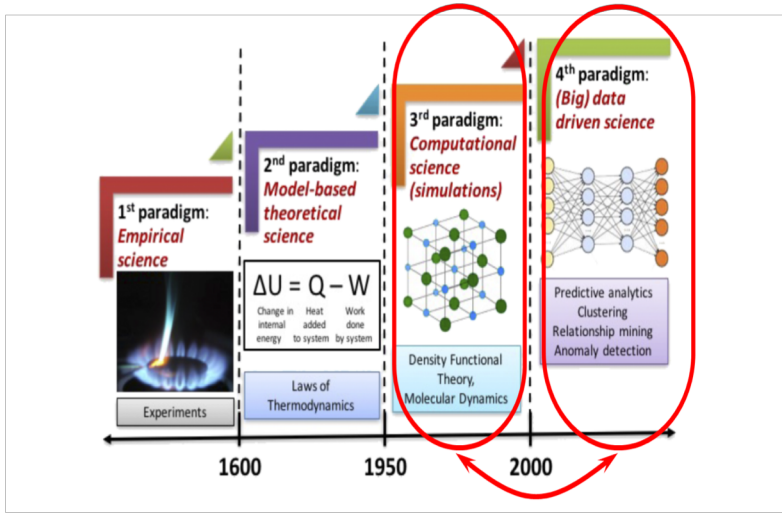
Dell PowerEdge R920
3 TB RAM

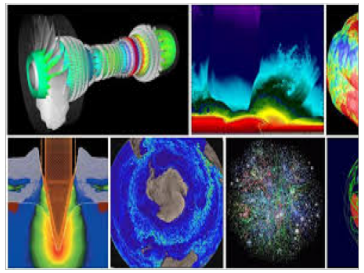
https://wiki.ncsa.illinois.edu/display/ISL20/Systems

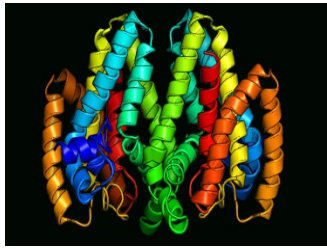
NCSA

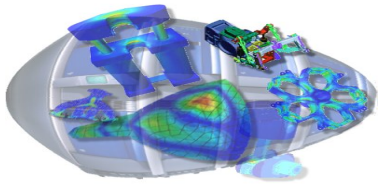
NCSA Industry Technical Team 2018 Highlights

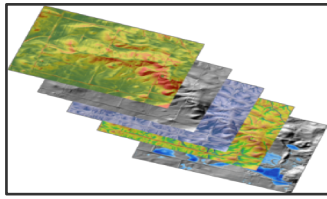
- Provide Technical Support and Consulting for the largest Industrial Supercomputing Outreach in the World
- 43 Industrial Partners (8 New in 2018)
- Domain Teams in Modeling and Simulation, Genomics and Data Analytics
- FTE Staff made of PhD-s and MSc-s, plus 21 students
- 33 Research Collaborators (15 UIUC faculty from 9 Depts.)
- 21 Publications
- Speak 8 world and 20 programming languages
- Work with hundreds of computing tools, libraries, codes and applications
- Center Directed Discretionary Award (CDDR)
- Best use of HPC in Automotive Award at SC 2018
- Key Note at ISC 2018 (Industrial Day)



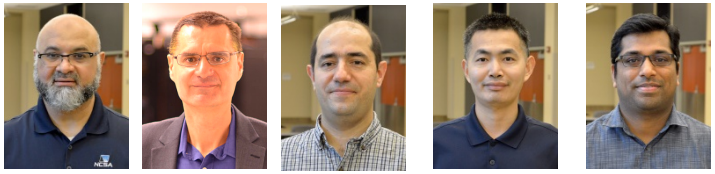








Modeling & Simulation "M&S" Team, NCSA Industry

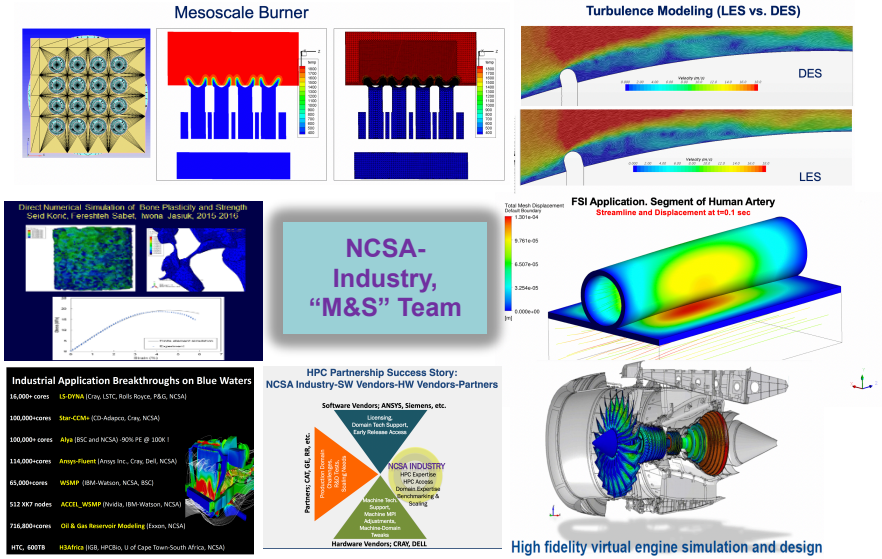


taha@illinois.edu

Updated January 29, 2019

Interests

- Domain Consulting & Applied Research
- Domain Application Benchmarking & Extreme Scaling
- Code Profiling, Parallelization, Performance Improvement & Optimization
- Deep Analysis of HPC Infrastructure and Applications
- HPC User Support & Training
- Strategic Collaborations Connecting Software and Hardware Vendors and Industry
- Public-Private Collaborations and Partnerships (local, regional, national and international)



Current Projects

- ***Modeling, Simulation and Fabrication of Bio-Hyprid Creatures*** – Living Organisms, Sensing, Actuation & Decision Making
- ***Convergence of Artificial Intelligence and Numerical Modeling in Turbulence*** – Deep Learning from Simulations to Accelerate Simulations
- **Multi-scale and direct numerical modeling of biomaterials and textile composites**
- **Multiphysics modeling of metal solidification processes**
- ***Improving Virtually Guided Certification for Product Design with Implicit FEA Software LS-Dyna***l – Parallel Scaling, Digital Twin, Regression Testing, Sparse Direct Solvers
- ***Simulation of Full Scale Gas Turbine***– 360 Degrees, All Cans, Turbine and Compressor
- ***Developing a Model for Predicting Engine knock Occurrence***– ICE, Engine Knocking Probability, Cycle-to-Cycle Variation
- ***Fluid-Structure Interaction "FSI": Tools & Applications*** – Fluid Flow, Structure Analysis, System Coupling
- ***CFD Analysis for High-Speed Electric Motors***– Fan Assemble, Blade Design, Cooling Load

Interest areas for collaboration/future work

Thermo-Fluid Applications Simulation - Combustion & Reacting Flow – Finite Elements & Structural Analysis - Fluid Structure Interaction "FSI" - Molecular Dynamics – Multiphysics & Multiscale - Materials Processing - Code Profiling, Parallelization, Modernization and Optimization – Sparse Direct Solvers - Extreme Scaling & Applications Benchmarking

Key Words:

Modeling & Simulation - Computational Fluid Dynamics - Finite Elements Analysis - Fluid Structure Interaction - Multiphysics, Multiscale - Molecular Dynamics - Nuclear Engineering - High Performance Computing – Code Profiling, Parallelization, Modernization and Optimization

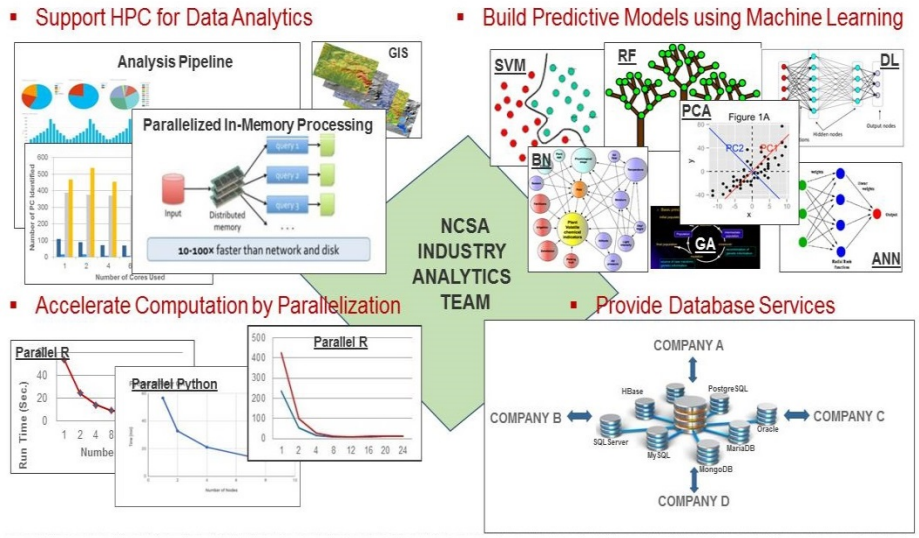
Data Analytics Group
NCSA Industry

doracai@illinois.edu

Updated January 18, 2019

Interests

- Data analysis in massive scale
- Real-world applications of Machine Learning
- Converge Artificial Intelligence with High Performance Computing
- Data management for very large databases and clouds
- Collaboration with industry partners
- High Performance Computing Support and Training



Current Projects

- ***A pharmaceutical company*** – Drug discovery, side effect analysis
- ***An agriculture company*** – Prediction on crop production, crop disease control
- ***An insurance company*** – Analysis on insurance policies, cost control
- ***An engineering company*** – Re-engineering workflows, simulating operations
- ***A manufactory company*** – Smart manufacturing, quality assurance, supply-demand control
- ***A finance company*** – Decision-making on loans, prediction on customer's credibility
- …

Interest areas for collaboration/future work

Applying Machine Learning on Industry problems; Applying High Performance Computing in problem solving; Application adoption, implementation, and modification; Code optimization; Process performance tuning; Data analysis and data management in massive scales; Geographic information System (GIS) for spatial data; existing and future workforce training to interest areas for collaboration/future work.

Key Words: Data Analytics, Machine Learning, High Performance Computing, Statistics, Predictive Models, Software Development, GIS, User Support, Consulting, Training

NCSA Genomics

Interests

Genomics, bioinformatics

- Variant calling
- Genome assembly
- Graph representation of reference genomes
- GWAS

Software, algorithms

- New software for massively parallel analyses on Spark
- Benchmarking software on different hardware
- Software performance optimization
- Machine learning approaches
- Workflows and automation for biological big data analyses

Infrastructure

- Cloud deployment of genomics workflows
- Containerization
- Production software testing for robust clinical deployment
- Developing best practices for genomics in HPC
- Genomics algorithms on FPGA

Current Projects

Mayo Grand Challenge

- Production genomics workflows for Mayo Clinic
- structural variants for hypoplastic left heart syndrome
- Performance optimization for genomic data compression

CCBGM

- Massively parallel implementations for epistatic GWAS
- Search space reduction for higher statistical power

CPRHD

- Computational Program for Racial Health Disparities
- Machine learning on large heterogeneous datasets

H3ABionet2.0

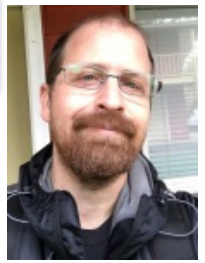
- Building a pan-African network of institutions with advanced expertise in computational genomics
- Developing approaches for graph-based reference genome representation, visualization and analysis
- Machine learning on biomedical data









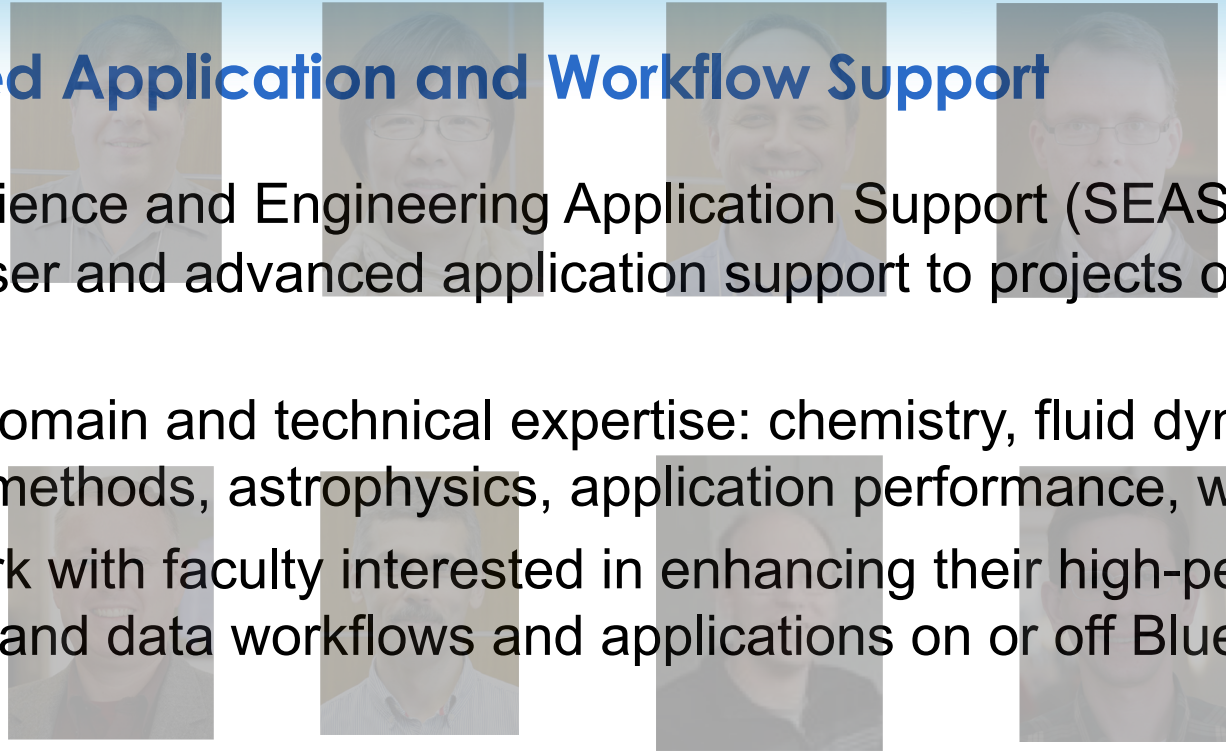



Interest areas for collaboration/future work https://wiki.ncsa.illinois.edu/display/LH/HPC+for+Computational+Genomics

Machine learning, racial health disparities, massively parallel analyses on Spark, workflow development, automation.

Key Words: Genomics, workflows, bioinformatics, high performance computing, big data.

Advanced Application and Workflow Support



- NCSA's Science and Engineering Application Support (SEAS) team provides user and advanced application support to projects on Blue Waters.
- Range of domain and technical expertise: chemistry, fluid dynamics, numerical methods, astrophysics, application performance, workflows.
- Able to work with faculty interested in enhancing their high-performance computing and data workflows and applications on or off Blue Waters.
- Gregory Bauer – SEAS team lead

Data Analysis and Visualization (DAV)

http://vis.ncsa.illinois.edu

Enabling scientific discovery through data-oriented research and development

- Support:
  - Software (VisIt, ParaView, IDL, ImageMagick, etc)
  - Data preparation
  - Best practices & training
- Research:
  - Data analysis (statistics, machine learning, etc)
  - “Is this in my data?”
  - “This is complex, can I show it?”
  - Vis for HPC
- Outreach: production-quality videos

NCSA