NSF COMPUTING FOR THE ENVIRONMENT

COMMITTEE MEMBERS & CONTRIBUTORS

*Magdalena Balazinska**
Terry V. Benzel
Azer Bestavros
*Vint Cerf**
*Andrew A Chien**
Gabriela Cruz Thompson
Mark E. Dean
Carla P. Gomes
*Roch Guerin**
Christopher R. Johnson
Raja Kushalnagar
*Ran Libeskind-Hadas**
Muriel Médard
Elizabeth Mynatt
Klara Nahrstedt
Timothy M. Pinkston
*Enrico Pontelli**

*Contributed to the development of the report.

National Science Foundation
WHERE DISCOVERIES BEGIN

CISE ADVISORY COMMITTEE MEMBERSHIP | DECEMBER 2022

Magdalena Balazinska

Advisory Committee Co-chair

Professor, Paul G. Allen School of Computer Science & Engineering Director, UW eScience Institute

Senior Data Science Fellow, UW eScience Institute

Director of the UW Advanced Data Science Option program, University of Washington

Terry V. Benzel

Director, Networking and Cybersecurity Division, Information Sciences Institute, University of Southern California

Azer Bestavros

Associate Provost for Computing & Data Sciences

Warren Distinguished Professor of Computer Science

Founding Director of the Hariri Institute for Computing, Boston University

Vint Cerf

Vice President and Chief Internet Evangelist Google

Andrew A. Chien

William Eckhardt Distinguished Service Professor, Computer Science

Director, CERES Center for Unstoppable Computing, University of Chicago

Gabriela Cruz Thompson

Director of University Research and Collaborations, Intel Corporation

Mark E. Dean

Emeritus Professor of Electrical Engineering and Computer Science, Tickle College of Engineering, University of Tennessee in Knoxville

Carla P. Gomes

Ronald C. and Antonia V. Nielsen Professor of Computing and Information Science

Director, Institute for Computational Sustainability Cornell University

Roch Guerin

Harold B. and Adelaide G. Welge Professor Department of Computer Science and Engineering Washington University in St. Louis

Christopher R. Johnson

Founding Director, Scientific Computing and Imaging Institute

Distinguished Professor, School of Computing, University of Utah

Raja Kushalnagar

Director, BS, Information Technology Program, and MS, Acces Human-Centered Computing

Professor, School of Science, Technology, Accessibility, Math and Public Health**,** Gallaudet University

Ran Libeskind-Hadas

Chair, Kravis Department of Integrated Sciences Claremont McKenna College

Amy McGovern

Lloyd G. and Joyce Austin Presidential Professor, School of Computer Science and School of Meteorology

Director, NSF AI Institute for Research on Trustworthy AI in Weather, Climate, and Coastal Oceanography, University of Oklahoma

CISE ADVISORY COMMITTEE MEMBERSHIP | CONT'D

Muriel Médard

Cecil H. Green Professor of Electrical Engineering Massachusetts Institute of Technology

Elizabeth Mynatt

Dean, Khoury College of Computer Science Northeastern University

Klara Nahrstedt

Ralph M. and Catherine V. Fisher Professor, Department of Computer Science University of Illinois at Urbana-Champaign

Timothy M. Pinkston

Professor and George Pfleger Chair in Electrical and Computer Engineering

Vice Dean for Faculty Affairs, USC Viterbi School of Engineering University of Southern California

Enrico Pontelli

Dean of Arts and Sciences New Mexico State University

Padma Raghavan

Vice Provost for Research

Professor of Computer Science and Computer Engineering Vanderbilt University

Rob A. Rutenbar

Senior Vice Chancellor for Research University of Pittsburgh

Yuanyuan Zhou

Qualcomm Chair Professor University of California, San Diego

SUMMARY OF RECOMMENDATIONS

The CISE community has much to offer to help tackle some of the greatest environmental challenges of our time. This document proposes a set of recommendations for the NSF and the CISE community to encourage and facilitate this type of research.

THRUST 1 | MULTIDISCIPLINARY COMMUNITY BUILDING

The first recommendation is to start with the identification of a few selected "Grand Challenges" that can serve as focal points around which respective, interdisciplinary communities can be built. The building of those communities can be initiated through the organization of Summits dedicated to each Grand Challenge. The organization of a summit should be assigned to a planning committee with representation from both CISE and domain experts.

THRUST 2 | OPPORTUNITY DISCOVERY

The second recommendation is to support new teams and junior researchers:

- Promote a model similar to NSF Planning Grants to bring multidisciplinary researchers together to develop shared research visions in computing for the environment.
- Target some graduate research fellowships specifically in this area, develop cluster fellowship programs to build graduate research communities, and offer support for graduate students and postdocs to visit the labs of active researchers in this field.

THRUST 3 | INFRASTRUCTURE BENCHMARKS/DATA/REUSABLE PRODUCTS

The third recommendation relates to creating resources to facilitate research:

- Invest in systematic creation of realistic environmental benchmarks and data sets for critical environmental challenge problems.
- Create cross-disciplinary collections of data and scientific models that capture current and dynamically evolving "state-of-the-art" understanding.
- Support workshops and hackweeks to encourage researchers to become more familiar with these kinds of resources.

THRUST 4 | PARTNERSHIPS

The final recommendation is to pursue cross-agency and private sector collaborations, as well as engaging with policy makers.

MOTIVATION

Environmental preservation is one of the greatest challenges of this generation. From climate change to loss of biodiversity, the challenges ahead of us are urgent and unprecedented.

The CISE research community has much to offer to help tackle those challenges, but it remains difficult for individual researchers to establish fruitful collaborations, identify problems that align with their expertise, and make an impact. Our overarching goal is to raise significantly the relevance of CISE in tackling today's environmental challenges by helping members of our community to more easily engage and contribute.

In this report, we identify some of the existing roadblocks and make concrete recommendations for the CISE community to create structures that will facilitate interdisciplinary collaborations and impact under the broad theme of "Computing for the Environment".

Our recommendations have four main thrusts:

1. In the "Multidisciplinary Community Building" thrust, we seek to create more solid and better integrated multidisciplinary communities, specifically under the Computing for the Environment umbrella. This includes identifying current hurdles for multidisciplinary collaborations and working on overcoming them.

2. In the "Opportunity Discovery" thrust, we seek to create mechanisms for the CISE community to identify 5-year and 10-year targets; specific problems that the CISE community can tackle based on existing environmental grand challenges.

3. In the "Benchmarks, Data, and Reusable Products Infrastructure" thrust, we aim to create infrastructures in the form of data, benchmarks, software, and hardware to help individual researchers plug in (discover, learn about, and understand the data and problems) and contribute toward the longer-term targets without the need to establish their own, multidisciplinary collaborations nor become an expert in the domain.

4. Finally, in the "Partnerships" thrust, we seek to initiate and strengthen the partnerships necessary for this initiative to succeed.

We also acknowledge that the CISE community can contribute along three distinct dimensions, all important:

- ***Improving the environmental footprint of computing systems.*** Those projects are usually CISE-centric, though not always, and target topics such as power management, computing architectures that are designed with reduced energy consumption as a primary goal, computing models and algorithms that include power consumption as an explicit variable or constraint, etc. This category includes CISE's current initiative in "design for sustainability". These core efforts are important, though their impact is limited in scope to the environmental footprint of computing systems, and their ability to go beyond their initial scope is often dependent on industry adoption, which is hard to realize.

- ***Using computing to decrease the negative environmental impact of large-scale human-made systems.*** This covers topics such as smart cities, smart buildings, power grid coordination, vehicle fleet management, environmental monitoring, precision agriculture, etc. There is often, though not always, a strong cyber-physical systems and Internet of things component to those projects as their ability to have an environmental impact usually calls for the combination of data acquisition and computation, that then controls some form of actuation to induce environmental improvements. These projects are inherently interdisciplinary with both computing and domain-specific expertise acting as peers, as both designing the right computing solutions calls for an in-depth understanding of the application area, and accurate modeling of complex dynamic interactions across domain, environmental, behavioral, economical, and social aspects. Today, successful deployments are relatively rare because it is inherently hard to account for these and many practical aspects that affect eventual adoption and successful application.

- ***Leveraging computing as a scientific discovery tool in areas with strong potential for environmental impact.*** ["Computational X for many values of X"] Today, this primarily encompasses the recent surge in artificial intelligence, machine learning, and reinforcement learning solutions for a wide range of fundamental problems that can have a major environmental impact. Examples include computational solutions that speed-up the search for new materials or molecules with specific physical properties that can drastically affect our environmental footprint, using deep learning systems to better control nuclear fusion, novel computer vision approaches to track animals in the wild using camera traps. Those projects are also deeply interdisciplinary, though the role of computing is often perceived as secondary, if only because the visible value of the end-product lies in a different domain with computing being mainly the "tool" that was used to get there. This perception can at times be an impediment to fostering the successful collaborations that such efforts require.

Through the set of recommendations below, we primarily target the last of the above dimensions, although perhaps the same recommendations could apply to all three dimensions.

KEY RECOMMENDATIONS

THRUST 1

MULTIDISCIPLINARY COMMUNITY BUILDING

The environmental challenges the CISE community can help tackle are inherently interdisciplinary, and interdisciplinary work is often impeded by external hurdles. These include differences in publication venues across communities, and possibly more importantly, a lack of awareness and understanding of the relevant data sets and ontologies that describe problems of importance. Overcoming those hurdles calls for a targeted effort to foster the development of interdisciplinary teams that bring together researchers from the CISE community and other scientific domains with relevant expertise, leading to effective communities of practice through collaboration. This is the goal of this first thrust.

A first step involves the identification of a few selected “Grand Challenges” that can serve as focal points around which respective communities can be built. The building of those communities can be initiated through the organization of respective Summits dedicated to each Grand Challenge. The organization of a summit should be assigned to a planning committee with representation from both CISE and domain experts.

The primary task of the planning committee is to structure the summit so as to address the difficulties commonly encountered in the formation of interdisciplinary collaborations involving CISE researchers (see also Thrust 3), inclusive of (1) Identification of foundational and accessible reading materials, (2) Selection of sample data sets, clarifying ontologies, and existing computational tools, and (3) Articulation of explicit goals and benchmarks to assess success.

The goals of the Summit are to enable CISE researchers to understand the Grand Challenge’s state of the art and familiarize themselves with relevant data and methodologies, and conversely introduce Domain Experts to the type of computational advances that may be developed towards addressing the Grand Challenge. Additionally, because of the social and societal implications associated with many environmental Grand Challenges, discussions at the Summit should also encompass aspects of policies, social acceptance, and how they may affect the development and deployment of solutions.

Summits should aim for broad participation from CISE and Domain researchers to enable the exploration of CISE-driven innovations, formation of interdisciplinary research teams, and exploration of both technological advances as well as policies and impacts on stakeholders.

It is expected that the activities of the interdisciplinary teams that will come out of the Summits will leverage and benefit from the support not only of the National Science Foundation (including the new NSF directorate on technology, innovation, and partnerships) but also other relevant aligned agencies and industry partners (see also Thrust 4).

As a concrete example of a grand challenge, we might ask whether our present societal focus on consumption and convenience could be altered beneficially. One-time use packaging and unrepairable products feed the throwaway economy that consumes energy and resources. Is there an opportunity to increase the societal value of products that can be repaired, repurposed, reused and recycled? Ironically, used cars increased in value over new ones in the recent pandemic because used cars were available immediately and new ones were delayed by supply chain disruptions. Rigorous study of such questions involves computational modeling of complex coupling of manufacturing, economics, circular recycling, and environment dynamics.

THRUST 2

OPPORTUNITY DISCOVERY

Many of the best ideas are likely to emerge by encouraging individual researchers to generate novel ideas and identify prospective collaborators. To that end, we promote a model similar to NSF Planning Grants to bring multidisciplinary researchers and institutions together to develop shared research visions in computing for the environment.

Computing for the environment is clearly compelling and attractive to young researchers and we recommend investment to nurture those interests. For example, CISE may wish to target some graduate research fellowships specifically in this area, develop cluster fellowship programs that help build graduate research communities in this space, and offer support for graduate students and postdocs to visit the labs of active researchers in this field. Additionally, undergraduate research programs in this area can be encouraged through NSF's REU program.

CISE can also play an important role in helping the community to develop, share, and explore datasets and benchmarks related to sustainability and the environment (as we discuss further in Thrust 3).

THRUST 3

INFRASTRUCTURE BENCHMARKS/DATA/REUSABLE PRODUCTS

Environmental and sustainability challenges often involve complex, interdisciplinary science spanning complex dynamic natural and human-made systems. Promising local changes in such settings often give rise to unexpected negative consequences. The complex internal behavior of such systems – atmosphere, ocean, biosphere, as well as human-engineered systems such as power grids, transportation networks, Internet, clouds, and more – and their coupling with each other implies that finding feasible, real, impactful solutions requires interdisciplinary pairing - data, problems, models, and of course scientific collaboration. For individual and small groups research efforts, this is difficult. To promote more impactful science on large-scale environmental and sustainability challenges by computing researchers, NSF CISE should:

- Invest in systematic creation of realistic environmental benchmarks and data sets for critical environmental challenge problems. These problems should be designed for coordinated multidisciplinary study, with challenges from the perspective of each relevant discipline, including accessibility to computing researchers. These benchmarks and data sets must be kept current, that is updated regularly to maximize the relevance and impact of the computing and other disciplinary research on the associated environmental challenges.

- Create cross-disciplinary collections of data and scientific models that capture current and dynamically evolving “state-of-the-art” understanding of the behavior of the environmental and human-engineered systems, as well as their coupling. Existing benchmark and data set approaches lack the ability to accurately model coupling across complex systems, rendering them unable to capture critical constraints and opportunities that are the key to realistic and high impact solutions. The proposed large-scale data and modeling infrastructures can accelerate scientific insight, and offer platforms for realistic evaluation that can inform public policy.

- Workshops and hackweeks could be supported to encourage researchers at all levels - from undergraduates to faculty - to become more familiar with these kinds of resources and gain experience trying to pose and answer questions using them.

NSF has historically engaged in efforts to accumulate and curate scientific data through its Research Data Alliance. It seems timely to explore how well this growing data collection is curated. Can it be mined for discovery? Can it be used to develop and validate models that inform our understanding of the environment and its complex dynamics?

The NSF NEON project and NSF Program Solicitation 21-549 for a “Center for Advancement and Synthesis of Open Environmental Data and Sciences” are also good examples of efforts in this direction.

THRUST 4

PARTNERSHIPS

NSF has a rich history of collaboration among the directorates of the Foundation, with other US and international agencies and, increasingly, with industry. The new TIP directorate enhances the opportunity to fashion broad and rich collaborations. The increasingly important role of computation, modeling, data analysis and machine learning to understand complex processes opens up a serious opportunity for CISE to promote multidisciplinary work in partnership with a wide range of players. NSF already engages with other departments in the US Government through the National Science and Technology Council, the Networking and Information Technology Research and Development (NITRD) Program and joint research projects with DARPA, DOE, NASA, NIST, USDA, and many other agencies. In addition, NSF's international office provides another avenue through which partnerships and projects can be explored with willing international partners.

The Summits activities described in Thrust 1 and the Planning Grants proposed in Thrust 2 should help create and strengthen partnerships. The infrastructure development activities in Thrust 3 should also be carried out in partnership with others. Such partnerships should include academic partners but, equally important, are partnerships with and transfers to private sector actors and purposeful engagement with domestic and international policy making bodies. The importance of engagement with policy makers cannot be over emphasized here, especially with regard to coping with environmental challenges. Informed policy making, bolstered with verifiable data and models, will play a key role in any successful response to our increasing environmental distress.

It might be proposed that the President's Committee of Advisors on Science and Technology (PCAST) be drawn into this discussion to reinforce the importance of cross-agency and private sector collaboration. NSF and its sister agencies have at their disposal both an extraordinary collection of diverse disciplinary expertise combined with some of the most powerful computing capability on the planet. These assets should be applied wisely and collaboratively to aid the effort to contain environmental degradation. The President might use PCAST findings to encourage cross-agency and industry cooperation.