



NATIONAL SCIENCE FOUNDATION  
2415 EISENHOWER AVENUE  
ALEXANDRIA, VIRGINIA 22314

**NSF 23-017**

## Dear Colleague Letter: Cyberinfrastructure, Data and Computation Opportunities for CMMI

---

October 21, 2022

Dear Colleague:

In recent years, developments in computing hardware, software, sensing and data storage and organization have combined to create a data-rich environment. The mathematical and computational tools to exploit and expand this environment are empowering our Nation to address great challenges and seize transformative opportunities. The scientific advances generated by these developments and their applications, jointly supported by the National Science Foundation's (NSF) Computer & Information Science & Engineering (CISE) and Engineering (ENG) Directorates, spur innovation, create jobs, strengthen security and preparedness, and improve quality of life for people across the country and around the world. While the core programs in ENG's Division of Civil, Mechanical & Manufacturing Innovation (CMMI) have actively supported the development of models and methods to exploit the new environment to address a variety of science and engineering problems, NSF has a number of opportunities that support the development and application of computational methods and data science to engineering problems as part of a broader cyberinfrastructure (CI) that will support both disciplinary and trans-disciplinary research communities. A comprehensive overview of NSF's vision for the future of cyberinfrastructure can be found at <https://www.nsf.gov/cise/oac/vision/blueprint-2019/Data-Software.pdf>. A working definition of cyberinfrastructure (Stewart et al. 2010) is "Cyberinfrastructure consists of computing systems, data storage systems, advanced instruments and data repositories, visualization environments, and people, all linked together by software and high performance networks to improve research productivity and enable breakthroughs not otherwise possible."

The focus of this Dear Colleague Letter (DCL) is to better acquaint the CMMI community with three of these funding opportunities, specifically:

- [Computational and Data-Enabled Science and Engineering \(CDS&E\)](#)
- [Cyberinfrastructure for Sustained Scientific Innovation \(CSSI\)](#), and
- [Training-based Workforce Development for Advanced Cyberinfrastructure](#)

(CyberTraining).

Each has distinct aims, requiring different issues to be addressed in a successful proposal. In the remainder of this DCL we shall briefly clarify the principal areas of interest of each program and offer some special points of attention for PIs to consider when responding to these funding opportunities.

## COMPUTATION AND DATA ENABLED SCIENCE AND ENGINEERING (CDS&E)

---

The goal of the CDS&E meta-program is to identify and capitalize on opportunities for major scientific and engineering breakthroughs through new computational and data-analysis approaches and best practices. It supports projects that harness computation and data to advance knowledge and accelerate discovery above and beyond the goals of the participating individual programs. For CMMI researchers, this means that a successful CDS&E proposal should, first of all, have the potential to make significant contribution to a research domain of interest to one or more of CMMI's core programs. Although not required, it is recommended that PIs identify the primary CMMI core program that best matches the proposed work.

**Proposals submitted to the CDS&E meta-program must also have a significant component of computational or data science that goes well beyond what would normally be included in these programs.** Thus, applying existing tools to problems of interest to CMMI's core programs is not sufficient for a successful CDS&E proposal; such a proposal should also demonstrate intellectual merit through enhancing computational and/or data science methods in a manner that will benefit research disciplines *beyond that in which they are developed*. A successful CDS&E proposal should clearly describe the specific nature of these methodological developments, making a case for how they advance the computational and data-driven methods of choice. Since the computational or data science is the critical component of the proposal, a clear validation and evaluation plan, describing the sources of data, amounts of data needed and generated, and the benchmarks against which the proposed approaches will be evaluated, is also important. Collaborative efforts with experimentalists are allowed; however, experimental components can only serve in a supporting role, e.g., providing/generating training or validation data. A discussion of how the proposed research will enable developments in scientific and engineering domains beyond the specific motivating problem will form an important part of the proposal's broader impact.

## CYBERINFRASTRUCTURE FOR SUSTAINED SCIENTIFIC INNOVATION (CSSI)

---

The CSSI program focuses on the development of data and software CI spanning the entire data and software lifecycle and the development, deployment and support of CI services for research communities ranging from small groups of individual investigators working on related problems to larger research communities. The emphasis in this program is the

development of CI services that enhance the productivity of a research community, supporting the domain-specific needs of science and engineering domains with tailored CI. While the CI would include enhanced computational and data-driven methods such as those that are the target of the CDS&E solicitation, a successful CSSI project should provide a robust service to a group of researchers addressing a demonstrated need in a manner that will enhance the group's ability to generate, evaluate and disseminate innovations within the group and to other communities who may be able to benefit. Thus, a successful CSSI proposal should have a clearly defined product, in the form of a CI service, with a clearly defined clientele and a clear description of the research bottlenecks it will alleviate if successfully developed. This latter involves a convincing argument as to why the proposed bottleneck cannot be addressed with existing commercial or academic tools. When submitting to the CSSI solicitation, members of the CMMI community should consider and clearly articulate the scientific challenges which the proposed CI would address, and how to engage the CMMI community as users.

### **TRAINING-BASED WORKFORCE DEVELOPMENT FOR ADVANCED CYBERINFRASTRUCTURE (CYBERTRAINING)**

---

The development, dissemination and deployment of CI hinges upon the availability of a diverse, skilled workforce ranging from researchers who are familiar with existing CI and able to apply it to their research problems to the hardware and software professionals implementing the data collection, communication and computing capabilities making up the CI. The CyberTraining solicitation seeks to address this issue from two complementary perspectives: by increasing the adoption of advanced CI by the research community, and by better integrating CI professionals into the research enterprise, facilitating the development of CI responsive to the needs of the research community. This program supports the development of innovative training and educational materials addressing emerging needs and identified bottlenecks in CI workforce development. When submitting to the CyberTraining solicitation, CMMI researchers should clearly identify workforce development outcomes.

### **SUMMARY**

---

In this DCL we have attempted to briefly summarize three mutually supporting funding opportunities in the area of cyberinfrastructure that should be of interest to CMMI researchers. Interested readers should make sure to read the specific funding opportunities listed above for each program carefully and should contact a cognizant program officer to clarify any questions before submitting a proposal. Finally, a unified presentation of the NSF's vision for CI can be found at [Transforming Science Through Cyberinfrastructure: NSF's Blueprint for a National Cyberinfrastructure Ecosystem for Science and Engineering in the 21st Century](#), which we believe will be of interest to many CMMI researchers.

A webinar discussing CMMI's participation in Data, Computation, and Cyberinfrastructure activities at NSF was held on Friday August 19, 2022, and the recording will be linked from the CMMI Website.

Sincerely,

Susan Margulies

---

## **References**

Stewart, C. A., Simms, S., Plale, B., Link, M., Hancock, D. Y., Fox, G. C., "What is Cyberinfrastructure?", SIGUCCS'10, October 24–27, 2010, Norfolk, Virginia, USA.