Coalition for Academic Scientific Computation

Sustainable Computing for Sustainability Workshop Response

20 NOVEMBER 2023

THEME CATEGORIES

Computing for Sustainability

- Identify a major sustainability problem for which computing may be able to help devise a solution (provide a general description).

We face many challenges today that will have profound negative impacts at a societal level if not immediately addressed. What's more, in some cases we lack solutions to these challenges altogether, or the current set of solutions are ineffective and unsustainable in varying degrees. Good examples of some of these consequential challenges include:

- The growth of economic inequality
- The paradox of social media being a cause of social isolation
- The impact on the environment from industrial, agricultural, mining and forestry activities
- The increase in human displacement due to socio-economic, environmental and political crises
- The ongoing risks of global pandemics to livestock and human health

A more hopeful view of our current predicament can be found in the fact that more sustainable and scalable solutions to the aforementioned challenges are being proposed and developed through the use of advanced computing and data-intensive approaches. As an organization of 100-plus member institutions involved in an incredibly broad spectrum of research, CASC members would be hard pressed to identify a single large-scale societal challenge we face today that is *not* in some way positively impacted by the use of advanced computing and data intensive approaches. In our experience, the use of computing helps to address these types of challenges in three ways. The first is simply in the identification of the problem. Examples of this include everything from socio-economic discrepancy in COVID-19 infection rates (https://doi.org/10.48550/arXiv.2205.04361) to combating disinformation in social media (https://doi.org/10.1016/j.tbench.2022.100035). While computing is a critically important tool in convincingly identifying problems, it also is used to develop solutions that

address these large-scale societal challenges. The literature is replete with examples of how computing has in some way led to unprecedented breakthroughs to address emerging challenges, from global pandemics to the development of new materials that sustainably store energy generated by carbon neutral sources. While the problems addressed by the use of computing have clear and tangible benefits to our society, computing is being used to solve more theoretical problems with profound impacts on the way we are able to address future challenges not currently known to us.

- Why is the problem important?
    - The problems that our member institutions help to address are important because they all, in some way, have a profound impact on the health and well being of our planet and its human population. More specifically, through the use of computing we have the opportunity to maximize human potential and minimize human suffering.
- How do you expect computing to address it?
    - Computing will be used to address future societal challenges in three general ways: 1) the identification of a problems before they have been widely acknowledged or discovered; 2) the development of effective solutions to address well understood but unanswered problems; and 3) the development of foundational methods to both identify and solve emerging problems. We now have many examples of comprehensive ways in which computing is coupled to research and discovery. For example, Artificial Intelligence (AI) approaches have for decades been proposed and developed without the major recognition of their current utility. Until the availability of massive data sets and suitable cyberinfrastructure becoming more generally available, we weren't able to fully realize the potential of this new approach to tackling large complex and contemporary problems. Likewise, we envision the use of increasingly detailed digital representations of objects (eg, digital twins) as a means of simulating solutions at many scales to explore the consequences of different approaches to address a broad range of societal challenges.
- Why is it not being addressed by existing initiatives?
    - Multidisciplinary approaches will be essential to solve many of the societal challenges that we face today. Traditional research solicitations tend to focus very narrowly on one or another aspect of the challenges that we face today. Furthermore, institutional processes can disincentivize large multi-disciplinary, multi-institutional and multi-sector collaborations. While single-contributor research and development will continue to be an essential part of our nation's research and development strengths, more attention needs to be paid to the promotion and funding of large-scale

multi-disciplinary, multi-institutional, and multi-sector proposals to ensure that the friction we currently experience between the development of an idea and its implementation is minimized. The difficulty of moving ideas into production is by itself a laudable area of study and certainly is among the most important reasons why challenges today are not being addressed by existing initiatives.

- What resources do you anticipate are needed to develop effective solutions?
    - The ability to plan, develop, and utilize cyberinfrastructure (CI) platforms and other resources depends not just on funding for technology. The primary enabler of the support services for computational research is the expertise of staff members for CI operations, consulting on hardware and software, data analysis and management, training, and direct participation in research project teams. Finding, hiring, and retaining people for these support positions requires recognition, credit, and respect for their work as part of the research enterprise, along with relevant job titles and career paths, steady funding, and changes in culture to allow a greater diversity of the workforce. In addition, flexible work environments are now a critical component to staff looking for their next opportunity. Almost every research area now requires access to advanced research CI resources, so skilled staff will become increasingly important members of the teams required to support cyber-enabled research efforts. Make no mistake, this need is reaching critical levels with the expanding need for CI resources across all areas of our nation's research and engineering enterprise and scholarly efforts. Research domains that previously did not engage in computational efforts now need easy-to-use resources and training to accomplish researchers' goals. As a result, the number of requests for assistance – and the need for more highly skilled CI staff - is rising quickly.

Sustainable Computing

- Identify a major sustainability challenge that computing is facing (provide a general description).
    a. While escalating energy requirements needed to operate cutting-edge data processing (CPU, GPUs, data storage elements and more) pose major challenges at local and national levels, our ability to effectively staff our core Research Computing and Data (RCD) facilities is perhaps our greatest sustainability challenge. CASC member institutions are facing challenges related to hiring staff. The issue is not how one institution fares against other institutions; it is about how academic RCD can overcome the issue and attract people from other organizations and industry. Hiring highly

experienced staff in RCD roles can be a nearly impossible undertaking, making developing and retaining existing staff a necessity.

- Why is the problem important?
    a. Staff members are the lifeblood and most essential piece of any organization, determining success or failure as much as – if not more than – any technology. Diversity within the staff, whether in life experience, domain expertise, or skill sets, enhances their contributions and ability to interface with diverse individuals.
- How is the problem unique to computing? (why would solutions from other fields not apply?)
    a. Research computing is complex, requiring above average IT skills and expertise far above those required for basic PC/Mac and Windows-based servers. Almost every research area now requires access to RCD resources, so an RCD team is a critical requirement to support cyber-enabled research efforts. There is an expanding need for RCD across all areas of campus, as research domains that previously did not engage in computational efforts now need easy-to-use resources and training to accomplish their goals. As a result, the number of requests for assistance – and the need for more RCD staff – has risen quickly.
- What dependencies exist to solving it? (social, political, economic, etc.)
    a. Support from University and HR: “This is a heavy lift and not for the faint of heart and requires patience. We have been working with our internal HR for 3 years now and are just now starting to see some positive results. The development of technical leaders & managers will be critical to helping address the problems both near and long term.” Our center directors need materials for use with university or organizational administration regarding the competitive nature of the HPC staff environment, justifying the unique job families and higher salaries required to recruit and maintain the staff necessary to support expanding research computing requirements. In some cases, user populations have doubled in the last 3-4 years. Developing close collaboration with high-level administrators such as VPRs and CIOs will ensure less duplication or friction among the various RCD staff positions. A formalized job description will help when working to define these RC-related positions. A better relationship with HR will hopefully bring the ability to recruit with diverse groups, both locally and globally. Changing hiring practices and modifying job descriptions should further increase the diversity of our talent pool. In addition, it has become clear that our centers and our industry overall should focus on diversity in recruiting, introducing younger students to the workforce – especially students in different disciplines – to advanced computing, and inviting them to experience activities that bring them closer

to computing careers. Sharing the materials from CaRCC and other groups mentioned above and in the “Resource Links” section at the end of this document with HR departments may help address this issue by providing examples of successful recruiting tactics.

b. Diversity in Recruiting: By now, most universities have produced a statement committing to diversity, equity, and inclusion in their hiring processes. These values must be emphasized in each job posting, and the university needs to follow through. Diversity is not just about statistics; it is also essential to ensure that diverse representation exists across all levels, including leadership. Centers would benefit from a recruitment push for more users from non-STEM domains who can work with faculty and show them what other researchers have done. However, finding staff who have that knowledge is a big hurdle. Student workers might have more domain knowledge, and we can help them with the computing knowledge. In addition, the ability to sponsor H1B/green cards would improve hiring rates with immigrants working at centers. In addition to pursuing broader dissemination of job postings, we need to prioritize access to RCD programs for HBCUs, small colleges, and other underrepresented groups. While there have been programs to bring underrepresented students to scientific research, many still lack an introduction to scientific research computing support, which could prove to be an interesting research career path.

- What resources do you anticipate are needed to develop effective solutions?

    a. Some of the actions that our member institutions are taking to develop effective solutions include:

        i. Improve Information Sharing for Jobs and Career Paths. CASC should post or link to example job descriptions and recommended career path options from CaRCC and other organizations to its website, where they can be shared with the community, and updated as needed.

        ii. Engage with and Promote a More Diverse Set of Institutions. CASC should look for ways to engage with HBCUs and other MSIs, including the National Association for Equal Opportunity in Higher Education (NAFEO) andthe Hispanic Association of Colleges and Universities.

        iii. Approach Workforce Development on a Community-Wide Basis. We are exploring ways for CASC Members to share their training and workforce development activities, so that staff from other centers can sign up or replicate

the training at their site. This could be done through the CASC website or a Google Calendar.

Building Interdisciplinary Teams

- What initiatives and resources would be most useful in helping bring CISE and "sustainability" researchers together?
    a. Incentivize the inclusion of CI professionals in Interdisciplinary teams. A greater variety of funding programs should include explicit provisions for CI professionals and workforce development as funded parts of their programs and solicitations. These should not be left to the institution to fund out of its own resources. The Federal Government should continue to expand the H1B visa program and other programs that allow hiring workers in specialized fields.
    b. Pursue Multi-Organization Partnerships. Federal research sponsors should continue to support programs such as Research Coordination Networks (RCNs), which seek to advance a field or create new directions in research and/or education by supporting groups of investigators to communicate and coordinate their research, training and educational activities across disciplinary, organizational, geographic and international boundaries.
- What major concern would you want to receive guidance on when it comes to putting together and overseeing interdisciplinary teams?
    a. A major concern of our members in the development and oversight of interdisciplinary teams is the lack of diversity among its members. In the best of worlds, end-to-end workflows bring together IT professionals, software engineers, molecular biologists, computational scientists, and mathematicians as a team with a common focus to find ways to leverage advanced detectors and instruments, fast networks, advanced computational platforms, and improved algorithms to produce entirely new insights and discoveries. These diverse teams working on every imaginable type of instrument are generating results that are fundamentally transforming the way we think about the world in fields like agriculture, medicine, and the environment. Well-run RCD resources are an essential component of Team Science, which is why this position is among our member's top concerns. Clear guidance as to the essential need for such individuals to be included as full-fledged members of these interdisciplinary teams is needed.
- What major challenge have you encountered in trying to build an interdisciplinary team?
    a. While progress has been made in the development of interdisciplinary teams, the challenges CASC members face persist. Developing solutions that lower barriers to integrating research computing and data resources more broadly within our campus and national research infrastructures is needed to be better equipped to support Team

science. These challenges fall under a broad set of areas, including cultural, technical, financial, operational, legal, regulatory, and strategic considerations.

b. Cultural and generational challenges start from the lack of common terminology, contributing to a lack of common awareness and understanding and to differences in prioritization among community members. These challenges can have negative effects on the timeliness of engagement and effective approaches to finding solutions. Technical considerations such as lack of common tools and/or difficulty in porting tools between environments aggravate these cultural challenges, leading to non-uniformity of user and developer tools. Security considerations often reduce opportunities to collaborate and develop common frameworks.
c. Growing legal and regulatory frameworks have accelerated the need for operational expertise in advancing cyberinfrastructure in Team Sciences. For research teams to navigate successfully through the complex issues of intellectual property, privacy and compliance, expertise is needed in the areas of researcher training, project management, and effective communications to stakeholders. There is no clear strategy that prioritizes RCD efforts, especially financially. This situation is aggravated by a general lack of planning and investment in research computing infrastructure, both on the part of funding agencies and the research community, as well as by a lack of reliable funding sources for long-term Team Science oriented RCD efforts.
d. Technical debt has also been a long-standing problem in the RCD arena that aggravates the assembly of effective tooling in team-based science. Issues include the lack of common tool sets in many areas, difficulty in porting legacy or familiar tools to new RCD ecosystems, and non-uniformity in skills for software development and use as well as for data management.
e. Financial and operational issues in Team Science include lack of funding or failure to fund development needed to support building and maintaining teams, lack of onboarding and training programs in certain areas causing a high threshold for learning for new participants, lack of coordination and communication among community participants and between these participants and the funding agencies, and a frequent lack of project management expertise in areas in which Team Science approaches are newly forming.