Title

The Early Years and Evolution of the DOE Computational Science Graduate Fellowship Program

Journal

Computing in Science & Engineering, 23(6)

ISSN

1521-9615

Authors

Brown, David
Hack, James
Voigt, Robert

Publication Date

2021

DOI

10.1109/mcse.2021.3120689

Peer reviewed

The Early Years and Evolution of the DOE Computational Science Graduate Fellowship Program

David Brown, *Lawrence Berkeley National Laboratory, Berkeley, CA, 94720, USA*
James Hack, *Oak Ridge National Laboratory, Oak Ridge, TN, 37830, USA*
Robert Voigt, *Leidos, Reston, VA, 20190, USA*

The U.S. Department of Energy Computational Graduate Fellowship Program, celebrating 30 years of existence in 2021, is one of the most successful graduate fellowships in the world as well as one of the longest running programs in the U.S. Department of Energy. This article discusses the conception, early years and evolution of the fellowship over the past thirty years.

FOUNDING THE FELLOWSHIP

When a group of scientists and engineers from across the U.S. gathered in a hotel meeting room in Washington, DC, in mid-October, 1990, they had little inkling that the new program they were founding was to become one of the most successful graduate fellowships in the world, as well as one of the longest running programs in the U.S. Department of Energy [DOE]. This committee met to develop a solution to a challenging problem for the U.S. DOE National Laboratories. Home to a good fraction of the world's high performance "supercomputers," the Laboratories were having a difficult time finding and recruiting new staff who had the skills to use those computers effectively to deal with the major interdisciplinary scientific and engineering challenges that the Laboratories were known for addressing. A general observation, agreed upon by those present, was that U.S. and international academia were largely unsuccessful in training graduate students in the skills needed to succeed in the newly emerging high-performance computing (HPC) world. A new four-year graduate fellowship program was proposed as a solution, and it needed to have some unique properties to successfully address the challenges. The attendees were well-known leaders in applied mathematics and high-performance computing from academia and government laboratories, including Peter Lax, Director of the Courant Institute, Robert Voigt, Director of ICASE, Paul Woodward, from the Army High Performance Computing Research Center in Minneapolis, Philip Colella, from the Mechanical Engineering Department at University of California at Berkeley, Susan Ying, from the Aeronautics Department at Florida State University, Paul Turinsky, from the Department of Nuclear Engineering at North Carolina State University, Ed Oliver from Oak Ridge National Laboratory, Patrick Burns from Colorado State University, James Hack from the National Center for Atmospheric Research, David Kuck from the University of Illinois, Edward Theil from Lawrence Berkeley National Lab, Jorge Moré from Argonne National Lab, Peter Jensen from Georgia Tech, and David Brown from Los Alamos National Lab. Representing the U.S. DOE Applied Mathematical Sciences (AMS) program were Gary Johnson and John Cavallini. The Oak Ridge Association of Universities (ORAU), who became the initial managers of the fellowship program, was represented by Craig Williamson.

It seems surprising now that one of the open questions was what to call the new program. The terminology used to describe the research activity was varied; one of the most common terms in use at the time was "scientific computing." Both "computational physics" and "computational fluid dynamics" (CFD) were well established by this time (albeit in a largely two-dimensional or "shallow water" world), but both terms described fields narrower than what was envisioned for this new program. The term "computational

1521-9615 © 2021 IEEE
Digital Object Identifier 10.1109/MCSE.2021.3120689

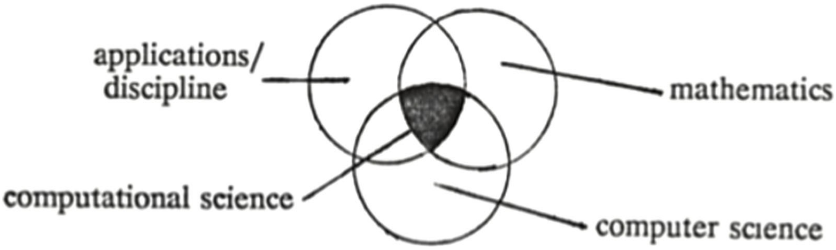

FIGURE 1. This Venn diagram describes computational science as the intersection of mathematics, computer science, and applied disciplines (*source: minutes of the 1990 CSGF advisory committee meeting*).

science" was not yet in common use but had the advantage over the other possibilities in that it clearly described a broad activity, and unlike "scientific computing," it implied that the research results were "science," not just perhaps ancillary "computing" results that supported science. In addition, the name naturally led to calling its practitioners "computational scientists." The committee members developed a working definition for the members of this new field: "a scientist or engineer who applies high-performance computational technology in an innovative and essential way to advance the state of knowledge in their discipline." And thus, the name for the new program: the Department of Energy Computational Science Graduate Fellowship. Notes from this meeting describe what was expected of the newly minted computational scientist: "The student is expected to design, implement, and analyze algorithms and computational techniques making use of the most advanced supercomputers and distributed computing environments available. Thus, the student must not only have a deep knowledge of an applied area but must also be proficient at using and understanding advanced computational tools. They are expected to be able to analyze methods and determine what will work best in which computing environments. As computing technology is so rapidly developing, they are expected to be flexible and forward-looking in their approaches."

A diagram sketched on the board was reproduced in the minutes of that first meeting (see Figure 1). It showed that computational science is made up of three parts: the applied discipline, mathematics, and computer science.

The key elements of the program were hashed out as well. One of the topics of discussion was which academic programs might be able to support students in the newly conceived program. Academic departments in such fields as applied mathematics, applied physics, geophysics, and computationally intensive engineering disciplines such as mechanical, aerospace, and chemical engineering, were recognized as potential homes for the students. The internet as we know it today was less than a decade old, and not all schools were connected to it, so the program would likely be limited to schools connected to a national network and with a local-area network in place. In the initial year of the program, a process was developed for the certification of universities that would be allowed to have fellows in the program. By the second year of the program, however, it had become apparent that certifying all participating universities was an impractical approach, and by the third year, the institutional certification process was dropped from the fellowship requirements altogether. Instead, the program focused on an instructional breadth requirement to assure that the students would receive the appropriate training in computational science. The student was required to submit an interdisciplinary "program of study" for approval that included courses in each of the applied mathematics, computer science, and science or engineering departments. By the 1998 application form, this had become a very specific requirement where at least one year of coursework in each of the three areas was required in the submitted program of study.

> *AVAILABILITY OF SUFFICIENT SUPERCOMPUTER TIME TO PERFORM THEIR RESEARCH WAS ONE ISSUE; ACCESS TO THE NETWORK FOR ACCESS TO THE SUPERCOMPUTERS WAS ANOTHER.*

In addition, a concern emerged about ensuring access for students to high-performance computing resources. Availability of sufficient supercomputer time to perform their research was one issue; access to the network for access to the supercomputers was another. Access to supercomputing time was to be provided by the National Energy Research Super Computer (NERSC) Center and other advanced computer research facilities. In the world of 1990, what we now call laptop computers were still an expensive novelty, and even access to a workstation that could connect to the network was recognized as a potential challenge for these new fellows. To address this, the program provided an allowance (to be matched by the student's academic department) of $2500 towards the purchase of a computer workstation. Beyond the single allocation for the workstation, a yearly institutional allowance of $1000 was to be provided to support the purchase of academic materials, travel to conferences, etc.

The committee members were also concerned about how to make this new program both visible and desirable. They recognized that one way to assure immediate interest in the program was to establish a premium stipend for the students. Four years of support was to be provided with a stipend starting at $1500 per month, increasing by $100 per month each year, and with an additional $300 per month to support expenses involved in relocation for a practicum opportunity (see description below). Adding full tuition coverage to the stipend support, this would place the new program among the most sought-after fellowships available at the time. For comparison, by 1998, U.S. National Science Foundation graduate fellowships still provided only $1200 per month for three years and capped the tuition support at $9500 per year.

The eligibility criteria for the program were to include:

1) entering graduate students;
2) students who had completed no more than one year of graduate study; and
3) students who had extensive graduate study but had not received departmental approval on a thesis topic.

Students who had a formally approved thesis topic would not be eligible for the program. Committee members also argued that the fellowship was unlikely to be able to influence students in the latter category to pursue a computational science research project of the kind envisioned by the committee. In later years, the eligibility criteria were changed to allow applications from undergraduate seniors and also restricted graduate applicants to those in the first year of graduate school.

EXPOSURE TO COMPUTATIONAL SCIENCE RESEARCH AT INSTITUTIONS SUCH AS THE NATIONAL LABORATORIES WOULD ALSO ENHANCE THE FELLOWS' EDUCATIONAL EXPERIENCE.

While coursework and academic infrastructure that supported a computational science experience for the fellows was a very important element of the new program, exposure to computational science research at institutions such as the national laboratories would also enhance the fellows' educational experience. Thus, a practicum experience at a DOE National Laboratory or equivalent was to be required for the fellows, which was to take place during one of the first two summers of the student's program. This was to be an opportunity for the student to be exposed to high-performance computing and participate first-hand in the interdisciplinary research that characterizes much of the activity at the national labs. Timing the practicum to occur early in the fellowship would have the most likelihood of influencing the student in their formulation of objectives for graduate school in this new research field of computational science.

Finally, the group discussed the desire to attract minority and women applicants to the program; ORAU implemented targeted advertising in minority and women's publications and committed to visiting HBCU's to promote the program. Since the beginning of the program, the CSGF advisory (now steering) committee has paid particular attention to the promotion of gender, racial and ethnic balance in the program, and statistics have steadily improved. Of the students admitted into the initial 1991 CSGF class, not quite a quarter were women. By the mid-2010s, the male/female balance had reached approximately 50%–50%, and by 2021, members of every racial and ethnic group tracked by the U.S. Department of Education were represented in the admitted class of fellows, a tribute to the success of outreach efforts to attract applications from traditionally under-represented groups.

The DOE AMS program, later renamed the Mathematics, Information, and Computational Sciences (MICS) Division, and now the Advanced Scientific Computing Research (ASCR) program, has supported the fellowship since its inception. Since 1999, the Advanced Simulation and Computing (ASC) program of the DOE National Nuclear Security Administration (NNSA) has also contributed to the financial support of the CSGF.

SELECTION OF THE FELLOWS AND MANAGEMENT OF THE PROGRAM

The selection process for choosing new fellows has evolved over the years. The first selection meeting, run by ORAU, occurred in the spring of 1991. Because of the hundreds of applications received that first year, probably due to rather nonspecific guidelines to potential applicants, it was held in a gym at Florida State University and engaged a very large number of reviewers from the computational science community and from a broad range of domain science and engineering fields. Reviewers spent two days selecting fellowship awardees and approving universities to participate in the program. Twenty-one new fellows at 15 universities formed the first class of the newly minted program.

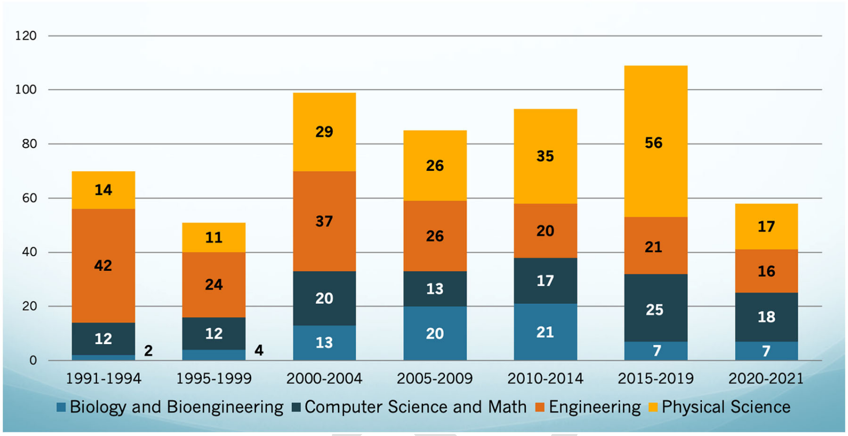

FIGURE 2. Disciplinary distribution of the CSGF over time (source: Krell Institute). Labels on the bars indicate the number of fellows in each category for the identified time interval.

In 1992, the management of the program shifted to the DOE Ames Laboratory under the leadership of James Corones, mathematics Professor at Iowa State University and Staff Member at Ames. Corones immediately introduced several changes in the program including a tightening of the application process, and the creation of a smaller, more active Advisory Committee consisting of five representatives from the computational science research community. He also named computer scientist Barbara Helland as Program Manager for the fellowship, a role she performed for over a decade before moving to DOE where she currently serves as the Associate Director for the Advanced Scientific Computing Research program. A selection process was established involving a prescreening that has evolved over the years to accommodate increasing numbers of applicants and a broader set of disciplines. The selection committee for the fellowship was made up of the Advisory Committee members plus three additional members from the computational science community. The improved specificity in the application resulted in 95 applications and 16 new fellows for 1993, the first year that Ames Lab managed the selection process.

In 1997, Corones founded the Krell Institute, a non-profit organization dedicated to supporting technology-based education and information programs, and management of the CSGF program moved to Krell. In 2003, the Advisory Committee was renamed the Steering Committee, recognizing the significant role it had played and continues to play in direction and oversight of the program. Over the years the Steering Committee has gradually increased in size to its present 11 members. Three members from the first Advisory Committee still serve on the present Steering Committee and four of its current members are alumni of the CSGF program.

Naturally, the selection process for the fellowship has also evolved, currently consisting of a multistage process in which a committee of 44 prescreens the hundreds of applicants each year, reducing the number to be considered in the final pool of applicants. Each application in the final pool is reviewed by every member of a 12-member selection committee with representation from national laboratories, academia, and private industry, many of whom served many years. These reviewers bring the wide spectrum of expertise needed to effectively evaluate the very large number of applications involving the broad disciplinary diversity that now exists in computational science. A large number of the current reviewers are alumni of the program.

While the founders of the program speculated that fellows would come from engineering, applied math, and applied physics programs, they also correctly envisioned that computational science would eventually pervade all scientific and technical disciplines. Indeed, over the years the applications for the fellowship have spanned a broad spectrum of science and engineering disciplines (see Figure 2). At the beginning of the program, a large fraction of what was then high-performance computation was focused on CFD or, slightly more broadly, computational physics. CFD research was primarily focused on aeronautics and astrophysics. Not surprisingly in the first couple of years most of the applications focused on these topics. Gradually more engineering applications appeared. Materials science became more popular whether the focus was understanding the material at the quantum level or at the macro level, for example, studying stress/strain relationships. By the early 2000s, interesting applications of computation to problems in biology and bioengineering such as modeling the behavior of proteins began to appear in the fellowship applications. As computing systems became more powerful and scientific experiments became more technically sophisticated, both simulations and experiments generated enormous amounts of data and fellowship applications focused on understanding how to make better use of data began to appear. Now with the emphasis on the roles of artificial intelligence and machine learning more and more applications are focused on how these tools can be used to discover patterns in data and to improve complex simulations. The emerging interest in quantum computing has led to applications looking at understanding such devices via quantum simulations and investigating algorithms that might be effective on such systems.

As part of their application, prospective fellows have always been required to describe a scientific or engineering project or area that they plan to pursue using computational science during their graduate studies. In 2018, in recognition that enabling technologies play an equally important role in computational science, the fellowship added a component focused on mathematics and computer science. Successful applicants must be receiving their degrees from mathematics or computer science departments and must be pursuing research that will contribute to advancements that will improve the effective use of high-performance computing. Although applicants do not have to focus on a specific scientific or engineering application, they are still required to include in their program of study courses outside of math and computer science that will expose them to scientific or engineering challenges where their enabling technologies might make a difference. The selection process for fellows in this component mirrors the careful procedure set out for the original CSGF program, currently employing a prescreening team of 10 and a final review committee of six representatives from the computational science community.

Over the years, the steering committee has worked with the management team at Ames Lab and later Krell Institute to enhance the experience of the fellows. A fellows meeting was instituted in 1993 and held biannually until 1999 when it became an annual meeting. This meeting, now called the CSGF Annual Review,[1] was conceived as an opportunity for the fellows to interact and better understand the program, to socialize within their peer group, for senior fellows to give presentations, and for program management and the advisory committee to hear about academic progress for the fellows and assess the overall progress of the CSGF program. With the exception of the first meeting, held in Minneapolis, and the 2000 meeting, held in the San Francisco Bay Area, these meetings have been held in the Washington D. C. area. This location has allowed much more participation by representatives of the sponsoring agencies, as well as opportunities for the fellows to interact with congressional representatives and staffers. In addition to the senior fellows' presentations, a poster session was added for the fellows in earlier years of the program, and also a poster session for the DOE Laboratories to present information about practicum opportunities. From 2009 through 2018, an HPC workshop was added to the Program Review agenda. This was replaced in 2019 by an opportunity for first-year fellows to attend the Supercomputing annual conference, which includes several days of tutorials and workshops on many aspects of high-performance computing.

DOE CSGF TODAY

The DOE Computational Science Graduate Fellowship Program has been vastly more successful than any of its founders could have imagined. Having started as a small initiative inside the DOE Applied Mathematical Sciences program, it has evolved into an internationally recognized and envied program now recognized as a line item in the U.S. Federal Budget. Over the past 15 years, from 349 to 729 students have applied annually for an average of 21 awarded fellowships. The Krell Institute, based in Ames, Iowa, continues to manage this fellowship under the leadership of Shelly Olsan, who replaced the late James Corones as its President in 2017. Krell now manages several fellowships for the U.S. DOE which are modeled after the CSGF. Since 2000, one or two of the CSGF graduating fellows have been recognized each year for their leadership, character, and technical achievement in the field of computational science by being awarded the Frederick A. Howes Scholar Award, named in memory of the DOE program Manager who guided and supported the fellowship in the mid to late 1990s. Nearly all the students in the program have spent at least several months at a DOE facility through participation in practicum assignments at all but five of the 21 DOE National Labs and Facilities (see Figure 3). (The 14 “other” practica were completed at other government and industry laboratories, which was permitted during the first few years of the program). Of the 431 living alumni of the program, Krell still tracks the career history for 418 (see Figure 4). The Fellowship has clearly been successful in its objective of providing a strong pipeline of talented computational scientists for DOE. More than a quarter of the fellows have held positions at one or more of the DOE Labs, and 63 are currently staff members at a DOE laboratory. A significant number of former fellows now hold leadership positions in the national labs, academia, and

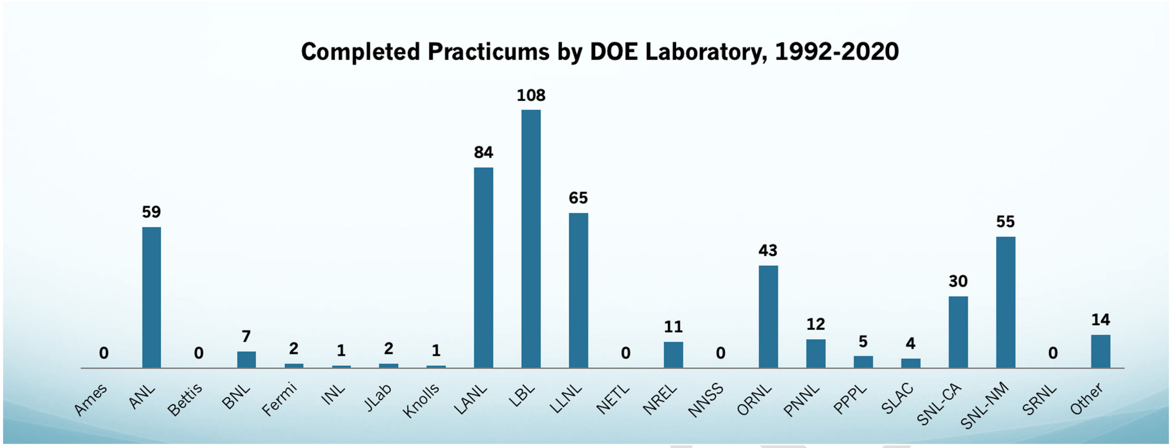

FIGURE 3. CSGF Practicum Assignment locations over the history of the program (source: Krell Institute).

industry, and have been recognized with numerous awards for their scientific achievements. The program has grown in size over the years, with 32 new fellows admitted in 2021, forming the largest class so far. Accomplishments of the fellows and alumni and progress of the program are captured annually in the DEIXIS magazine, which can be accessed online through the Krell website.[2]

The CSGF Alumni Association was formed and incorporated as a 501(c)(6) organization to join former DOE CSGF fellows together in organized efforts to benefit members of the Association and current fellows, to advocate for the DOE CSGF program, and to advocate for computational science and the computational science community.[3] The fellowship program has successfully adapted to the continual developments in computational science and will clearly continue to do so for many years to come. The modern international community of computational scientists is clearly indebted to the vision and creativity of that committee of scientists and engineers who, 31 years ago, understood the potential for computational science as an important element of the scientific enterprise, and conceived of and designed the DOE Computational Science Graduate Fellowship program.

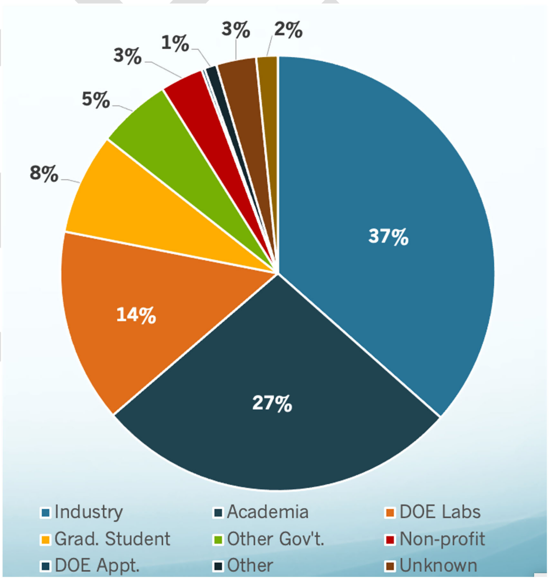

FIGURE 4. Employment status of CSGF Alumni (source: Krell Institute).

ACKNOWLEDGMENTS

The authors would like to thank Shelly Olsan and Lisa Frerichs of the Krell Institute for providing historical and statistical information about the CSGF. This work was supported in part by the U.S. Department of Energy, Advanced Scientific Computing Research Program under Contract DE-AC02-05CH11231.

REFERENCES

1. [Online]. Available: https://www.krellinst.org/csgf/conf
2. [Online]. Available: https://www.krellinst.org/csgf/outreach/deixismagazine
3. [Online]. Available: http://csgf-alumni-association.org/

DAVID BROWN is an Applied Mathematician and the Director of the Applied Mathematics and Computational Research Division, Lawrence Berkeley National Laboratory, Berkeley, CA, USA. He first became involved with CSGF as a Staff Scientist with Los Alamos National Laboratory, Los Alamos, NM, USA. He was also a Scientist with Lawrence Livermore National Laboratory from 1998 to 2011. He was a member of the CSGF advisory and steering committees from 1990 until 2020, and has served many times on the CSGF review committees for the program. He is currently co-PI of the CSGF program. He is the corresponding author of this article. Contact him at dlb@lbl.gov.

JAMES HACK is a Computational Climate Scientist and former Director of the National Center for Computational Sciences with the Oak Ridge National Laboratory. Prior to joining Oak Ridge, he spent 23 years at the National Center for Atmospheric Research, where he first became associated with the CSGF. He served on the CSGF advisory, steering and review committees, and currently chairs the steering committee. Contact him at jhack.ornl@gmail.com.

ROBERT VOIGT is a Computational Mathematician who has served in high-performance computing positions with the National Science Foundation and the Department of Energy. He began his association with the CSGF in 1990 while Director of ICASE. Since then, he has served continuously on the CSGF advisory, steering, and selection committees. He is currently a Consultant with the National Nuclear Security Administration's Predictive Science Academic Alliance Program. Contact him at rvoigt@krellinst.org.