ASCAC Subcommittee on Workforce Development Needs

Prepared by

Dr. Barbara Chapman
University of Houston
Subcommittee Chair

DOE ASCAC Subcommittee Report
November 21, 2014

CACDS

The Charge

identify disciplines that need more workforce training at grad student or postdoc level for DOE Office of Science mission needs.

Things to consider:

- Disciplines not well represented in academic curricula;
- Disciplines in high demand, nationally and/or internationally, resulting in difficulties in recruitment and retention;
- Disciplines where the DOE labs may play a role in providing needed workforce development;
- Specific recommendations for programs at the graduate student or postdoc levels to address discipline-specific workforce development needs.

Letter report on findings and recommendations by June 30

The Subcommittee

Name	Affiliation
Barbara Chapman	University of Houston (ASCAC) (Chair)
Henri Calandra	Total SA
Silvia Crivelli	Lawrence Berkeley National Lab and University of California Davis
Jack Dongarra	University of Tennessee (ASCAC)
Jeffrey Hittinger	Lawrence Livermore National Lab
Chris Johnson	University of Utah
Scott A. Lathrop	NCSA, University of Illinois Urbana-Champaign
Vivek Sarkar	Rice University (ASCAC)
Eric Stahlberg	Advanced Biomedical Computing Center
Jeffrey S. Vetter	Oak Ridge National Laboratory
Dean Williams	Lawrence Livermore National Lab (ASCAC)

Interpretation of Charge

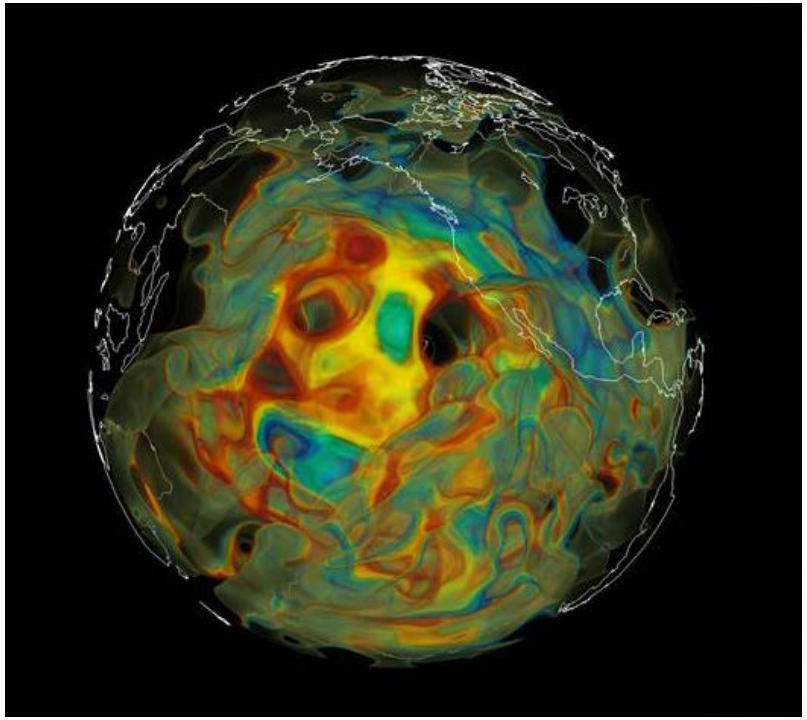
- ASCAC disciplines needed for DOE mission do not fit neatly into traditional academic categories
  - Some positions require multidisciplinary training
- We coined term **Computing Sciences** for this report
  - Includes: Algorithms; Applied Mathematics; Data Analysis, Management and Visualization; Cybersecurity; Software Engineering and High Performance Software Environments; and High Performance Computer Systems.
- Committee decided to take a broader perspective
  - in order to address workforce challenges identified, more than a program is needed

Limitations of Report

- Information on the disciplines (and multidisciplinary expertise) needed to support DOE mission is not readily available
  - No clear fit within traditional boundaries of academic disciplines
- Data and information obtained usually covers some part of the Computing Sciences
- HR data typically not available for this range of subjects
  - E.g. some positions are categorized in area of domain expertise
- Hence our findings cannot be conclusive

Role of Advanced Scientific Computing at DOE/SC

- Scientific computation is at core of much of Office of Science R&D.
- ASCR facilities, experts in their utilization are essential
  - To all areas of scientific activity in the DOE national laboratories.
  - Physics, chemistry, astrophysics, energy,…
- ASCR provides critical technology that enables DOE's world leadership in scientific computation



CACDS

Need for Skilled DOE Workforce in Computing Sciences

- Breadth of expertise required to effectively deploy tools increasing.
- Maintaining a sufficient workforce is this area is critical
  - To enable research outcomes
  - Amortize significant investment in ASCR facilities

Workforce Recruitment/Retention Challenges

- Labs invited to provide information on areas where they experience recruitment / retention difficulties
- Areas in Computing Sciences figured prominently in responses, especially in large labs
- For Computing Sciences, labs reported:
  - Low number of qualified applicants, many of which are foreign nationals
  - Takes a long time to fill positions
  - Labs spend significant effort on recruiting in this area

- Additional information solicited from larger labs to quantify problem
  - No uniform method for recording information; not all labs provided data
  - Takes labs twice as long as industry to fill positions in computing sciences
  - About 4 times as long when security clearance does not permit foreigner
  - Acceptance and retention rates mostly favorable

Competencies	National Laboratories
Advanced computing architectures	LLNL, ORNL
Applied mathematics (including advanced modeling and methods)	ANL, LBNL, LLNL, ORNL
Computational sciences/simulation; scientific software	ANL, BNL, INL, LBNL, LLNL, ORNL, SNL
Cyber security	INL, LLNL, ORNL, SNL
Data acquisition software	FNAL, ORNL
Data informatics (data mining, machine learning, big data, statistical techniques)	ANL, LBNL, LLNL, ORNL, PNNL
Dynamic mesh algorithms	LLNL
HPC /extreme-scale/exascale computing	ANL, INL, LANL, LBNL, LLNL, ORNL, PNNL, SNL
Performance analysis of HPC applications	LBNL, ORNL
Software quality assurance	LLNL, ORNL
Solvers	LBNL, LLNL
Storage systems	LBNL, ORNL
Uncertainty quantification	LLNL, ORNL
Visualization and scientific data analysis	LLNL, ORNL

Recruitment/Retention at Labs

Lab	# Open Posns	Ave. Time To Fill (Days)	Total Technical Staff	% Foreign Nationals	Declined Job Offers	Attrition Rate (%)
LANL	148*	263*	1903*	5.4*	21/173*	4.9*
LBNL[1]	56	112	206	38.4	2/39	8.0
LLNL[2]	146	311	2094*	7.4*	7/36	4.8*
ORNL[3]	87	110	379	38	11/73	7.6
PNNL[4]	44	107	1113*	16*	16/50	8.9**

* Data for all scientific and engineering disciplines, M.S. and Ph.D. level

** Data for all scientific and engineering disciplines, all degree levels

[1] LBNL data for "all scientists and engineers on the Computer Science curve"

[2] LLNL data based on best attempt to identify positions in the Computing Sciences; time-to-fill may be skewed by indefinite postings; attrition rate corrected for voluntary separation program

[3] ORNL data for "lab-wide computing/computational science" positions; attrition rate corrected to account for voluntary separation program (37% of terminations)

[4] PNNL attrition rate is uncorrected for voluntary separation program; historical rate is 4-5%; total number of job offers is estimated.

Workforce Recruitment/Retention Challenges

- Strong demand for M.Sc, Ph.D.-level Computing Sciences positions
  - Especially at LBNL and ORNL where open positions are ca. 25% of total staff in Computing Sciences
- Labs take longer than industry to fill positions in Computing Sciences
  - On average ca. 100 days to fill an Office of Science position
  - More than 200 days for NNSA lab positions
  - Industry needs 39 days on average for Computer and Mathematical occupations (all degrees)
  - 48-50 days on average for M.Sc., Ph.D. positions in STEM
- Attrition rates compare favorably with industry
  - 10% in industry, 5% NNSA labs, ca. 8% SC labs
  - Warrants further study, since loss of expertise can be catastrophic

FINDING: The multidisciplinary national labs face workforce recruiting and retention challenges in Computing Sciences

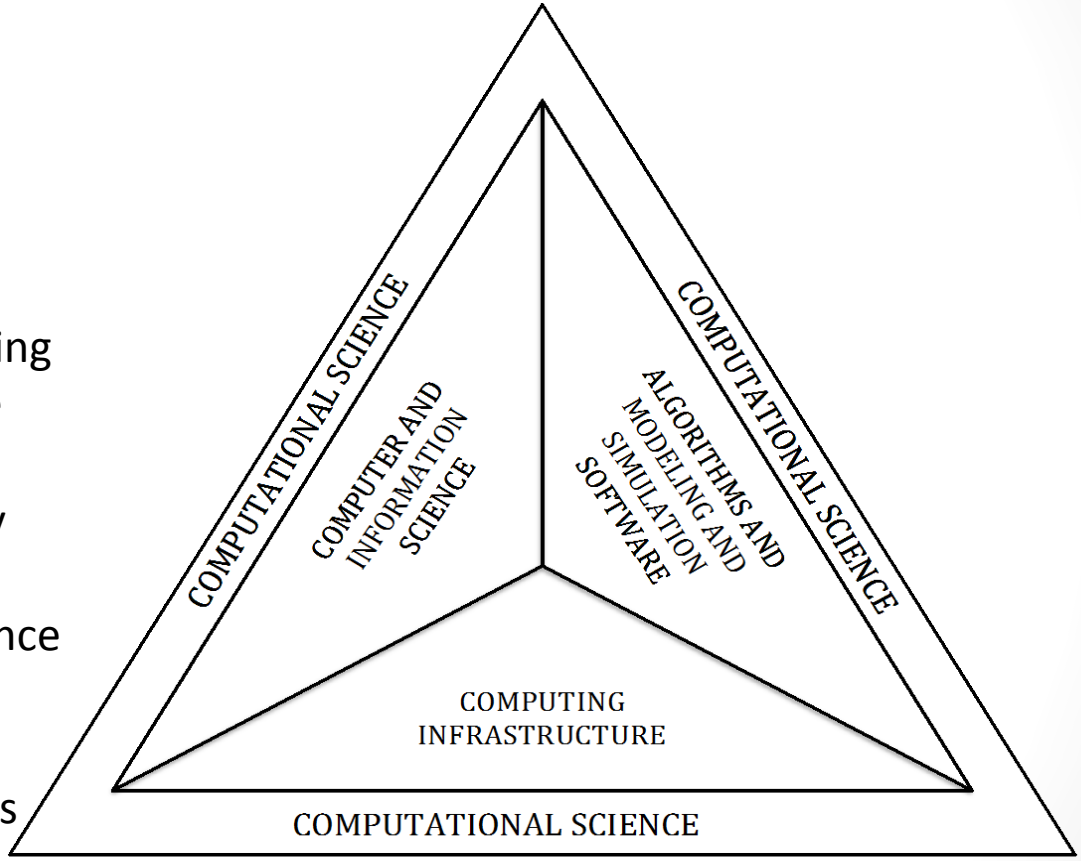
Multidisciplinary Education

Computational Science disciplines taught in:

- Computer Science
- Computer Engineering
- Information Science

Also in interdisciplinary studies:

- Computational Science and Engineering
- 2005 PITAC report described difficulties establishing these



From PITAC Report, 2005

CACDS

Multidisciplinary Education

- Interdisciplinary Computational Science and Engineering (CS&E) studies emerging
  - Domain sciences, applied mathematics, numerical analysis, computer science
  - Problem-solving methodologies, science and engineering tools
  - Degree program, an area of specialization or a certificate
- Cannot impart complexities of field, do not provide full skillset needed by DOE
- Insufficient number of graduates

- http://www.siam.org/students/resources/report.php has SIAM's 2014 listing of CS&E programs

Number of PhDs in CS and CE

PhD Specialty	2010	2011	2012	2013	Total
Artificial Intelligence	181	193	203	171	748
Databases/Info Retrieval	99	106	122	125	452
Graphics/Vis	87	111	99	99	396
HW/Architecture	78	70	92	91	329
High Performance Computing	29	37	49	60	175
Networks	150	140	147	152	589
Operating Systems	59	55	66	55	235
Scientific/Numerical Computing	33	27	32	29	121
Software Engineering	126	147	149	140	562

Taulbee Survey 2014

Disciplines in Academia

- Fewer graduates in fields related to Computing Sciences in Computer Science (CS), Info Science, Computer Engineering (CE)
  - Than in traditional areas
- NSF taskforce (2011): Universities not teaching essential skills for applying CS&E in the field, not preparing students to harness powerful new supercomputing
- Interdisciplinary Data Science education beginning to emerge, but not likely to satisfy demand

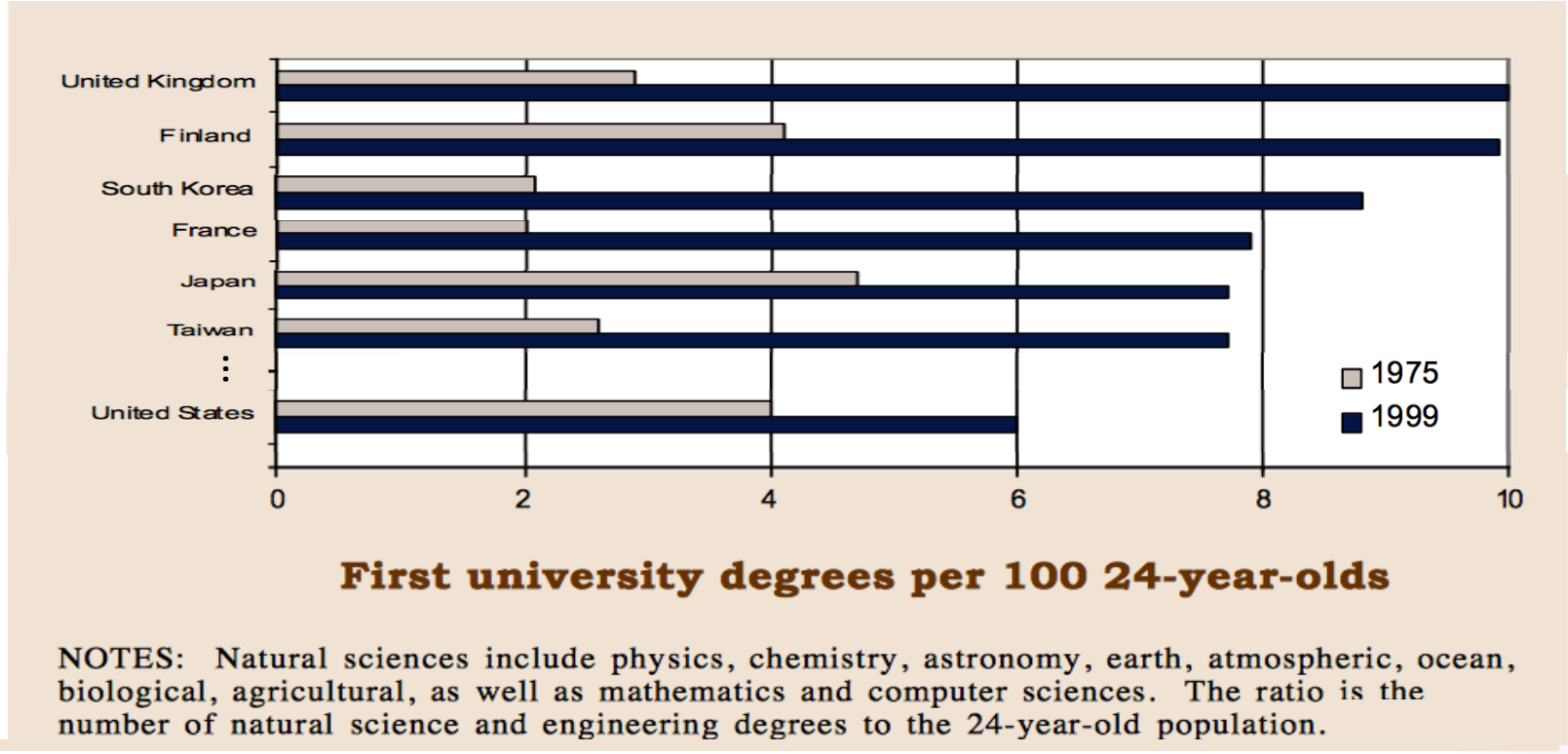
FINDING: Insufficient educational opportunities are available at academic institutions in areas of the Computing Sciences most relevant to the DOE mission.

Workforce Demand

- Huge demand for graduates with computing expertise
  - Taulbee survey reports very low unemployment in US for computing graduates (under 1%)
  - Large majority of graduates enter industry (70%)
- Retirement of current workforce is expected to grow workforce gap over coming decade
- Labs cannot compete with industry compensation
- Awareness of lab careers among graduates low
- Conference travel restrictions impede recruitment
  - Also decrease attractiveness of jobs

Council on Competitiveness: HPC is a proven game-changing technology

Ratio of NS&E First University Degrees to 24-year-old Population, 1975 and 1999



NOTES: Natural sciences include physics, chemistry, astronomy, earth, atmospheric, ocean, biological, agricultural, as well as mathematics and computer sciences. The ratio is the number of natural science and engineering degrees to the 24-year-old population.

National Science Board: The Science and Engineering Workforce. Realizing America's Potential. August 2003

An Incomplete Talent Pool

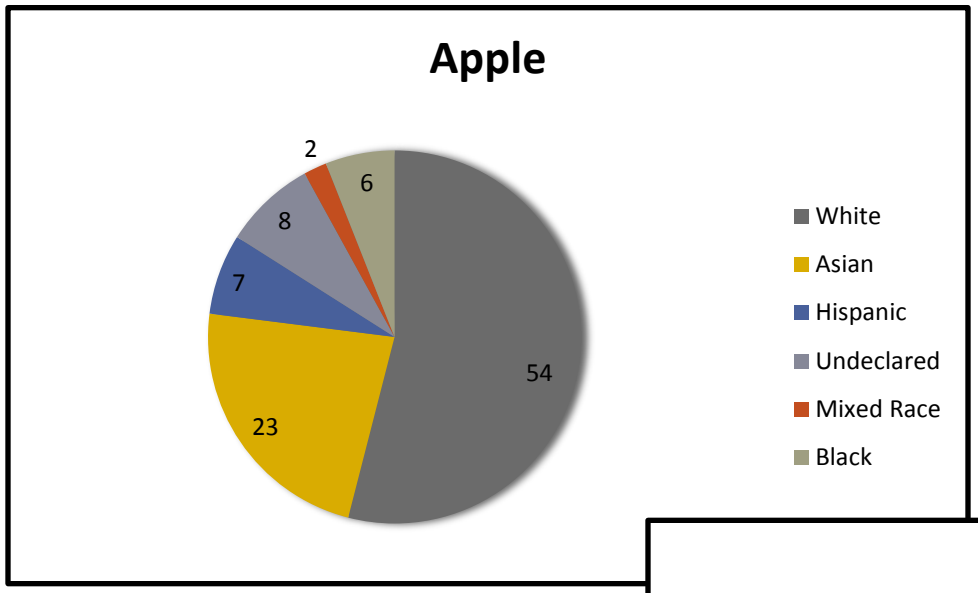
Lack of diversity in US graduates in CS and CE is a major contributing factor in national shortage:

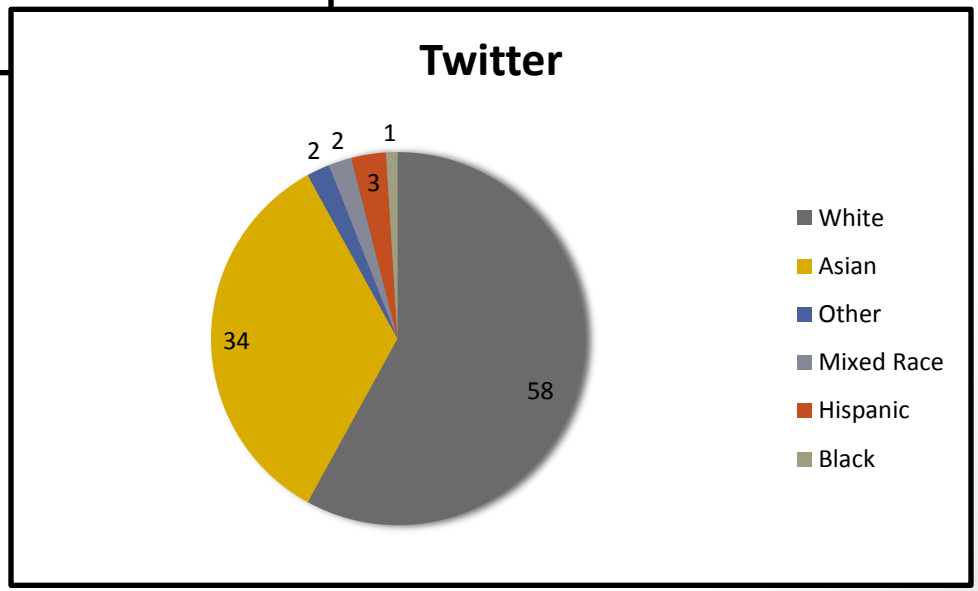
- US citizens among graduates are mostly white male
- African American / Hispanic graduates very low (ca. 1% each)
- Percentage of females among graduates is declining

Race/ ethnicity and sex	Total 24-year-old Population	Total bachelor's degrees	Total NS&E degrees	Bachelor's degrees per 100 24-year-olds	NS&E % of bachelor's	NS&E degrees per 100 24-year olds
Total	3,403,039	1,199,579	205,355	35.3	17.1	6.0
Sex						
Male	1,714,571	525,714	128,481	30.7	24.4	7.5
Female	1,688,468	673,865	76,874	39.9	11.4	4.6
Race/ethnicity						
White	2,251,292	878,018	142,500	39.0	16.2	6.3
Asian/Pacific Islander	149,413	69,988	22,003	46.8	31.4	14.7
Underrepresented minority	1,002,334	181,709	25,820	18.1	14.2	2.6
Black	473,402	95,878	12,731	20.3	13.3	2.7
Hispanic	497,620	78,125	12,006	15.7	15.4	2.4
American Indian/Alaskan Native	31,312	7,706	1,083	24.6	14.1	3.5

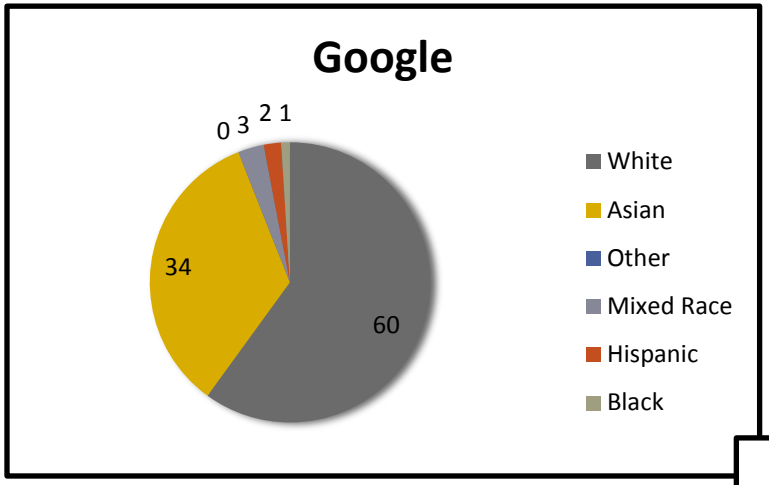
Participation Rate in Natural Sciences and Engineering Bachelor's Degrees in 1998

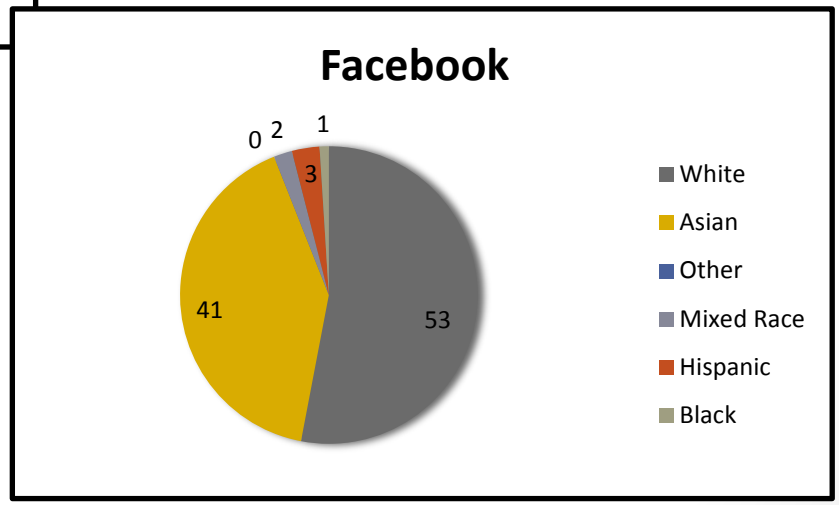
Demographic Trends in Industry





Demographic Trends in Industry





CACDS

US Citizens / Permanent Residents as Percentage of PhDs in CS, CE Areas

PhD Specialty	Citizens, Permanent Residents	% of Total
Artificial Intelligence	439	58.7%
Databases/Info Retrieval	203	44.9%
Graphics/Vis	228	57.6%
HW/Architecture	147	44.7%
High Performance Computing	78	44.6%
Networks	205	34.8%
Operating Systems	108	46.0%
Scientific/Numerical Computing	78	64.5%
Software Engineering	328	58.4%

Figures accumulated over past 4 years; Taulbee Report, 2014

LBNL Demographics in STEM

Types of Jobs at Berkeley Labs	TTL	Women	%	URM	%	OPC	%
Scientists and Engineers (Conducting Research)	640	100	15.6%	29	4.5%	131	20.5%
Postdoctoral Scientists	486	133	27.4%	26	5.3%	209	43.0%
Engineers (Information, Mechanical, and Electrical)	483	102	21.1%	51	10.6%	118	24.4%
Research Support (Non S&Es in programmatic divisions)	907	390	43.0%	145	16.0%	207	22.8%
Ops Support (Non S&Es in Operational Divisions)	677	324	47.9%	161	23.8%	117	17.3%
Totals	3193	1049	32.9%	412	12.9%	782	24.5%

Date of Data: October 1, 2013
Career, term, and postdoctoral employees only
URM=Underrepresented Minorities (African American/Black, Hispanic/Latino, and American Indian/Ala
OPC=Other People of Color (Asian/Asian American, Middle Eastern/Southwest Asian/North African anc

Data from DOE labs reflect national demographics

Also indicate retention problem for female postdocs

CACDS

Workforce Demand

- % of foreign nationals in graduate population growing steadily
  - 58% of graduates in Computing Sciences are now foreign nationals
- Lack of diversity in US graduates in CS and CE is a major contributing factor in national shortage
  - Current US citizens among graduates are mostly white male
  - African American / Hispanic graduates very low (ca. 1% each)
  - Percentage of females among graduates is declining
- Data from DOE labs reflect national demographics
  - Also indicate retention problem for female postdocs
- Lack of STEM diversity widely acknowledged, not addressed

FINDING: There is a growing demand for graduates in Computing Sciences that far exceeds the supply. A larger workforce gap and continued underrepresentation of minorities and females are expected.

DOE Workforce Development
Office of Science

Workforce Development Training Program

- Science Undergraduate Laboratory Internships (SULI)
- Community College Internships (CCI)
- DOE Office of Science Graduate Fellowship (SCGSF) Program
- Albert Einstein Distinguished Educator Fellowship (AEF) Program
- Visiting Faculty Program (VFP) at DOE Laboratories
- DOE National Science Bowl (NSB)
- STEM Resources for K-12 Educators

See http://science.energy.gov/wdts

Existing Workforce Training

- DOE CSGF program established 1991; jointly funded by ASCR and NNSA
- Trains graduate students to meet national workforce needs in computational sciences, including those of DOE
- Provides practical work experiences at DOE labs; improves collaboration between labs and academia; raises visibility of careers in computational sciences
- Effective elements
  - Interdisciplinary program of study
  - Research practicum at DOE laboratories
  - Annual review that enables networking
  - Careful selection process

CSGF Applicants' Quantitative Measures for 2002-2011

Year	Average UGPA	Avg. Percentile GRE Verbal	Avg. Percentile GRE Quantitative
2002	3.51	74	87
2003	3.60	75	85
2004	3.62	75	82
2005	3.59	73	83
2006	3.61	75	82
2007	3.68	75	85
2008	3.64	78	87
2009	3.60	79	86
2010	3.59	77	84
2011	3.64	77	85

CSGF Awardee Quantitative Measures for 2002-2011

Year	Average UGPA	Avg. Percentile GRE Verbal	Avg. Percentile GRE Quantitative
2002	3.72	77	90
2003	3.86	86	90
2004	3.90	83	88
2005	3.73	80	88
2006	3.92	85	89
2007	3.87	86	89
2008	3.80	86	91
2009	3.86	85	92
2010	3.81	87	91
2011	3.88	90	90

Breakdown of CSGF Applicants by Major Field of Study for Years 2002-2011

Area	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011
Bio & Bioeng	13	40	43	49	66	72	64	60	80	108
Math & CS	41	90	76	74	108	71	69	76	113	115
Engineering	65	113	116	133	149	150	152	125	194	243
Physical Sci	27	58	74	71	69	92	76	73	119	134
Social Sci	1	1	1	1	3	1	1	1	3	1
Did Not Report	12	16	12	10	15	10	9	14	21	27
Total	159	318	322	338	410	396	371	349	530	628

Breakdown of CSGF Awardees by Major Field of Expertise for Years 2002-2011

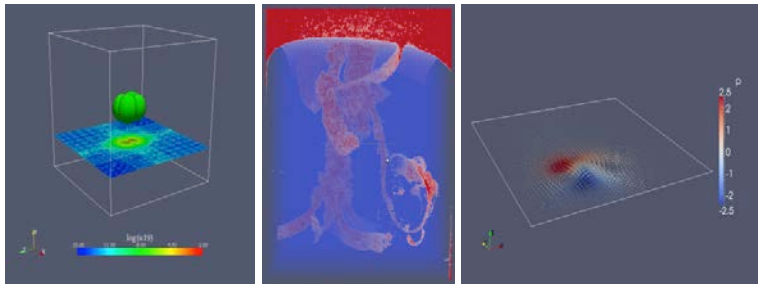
Area	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011
Bio & Bioeng	3	1	1	3	3	5	6	3	2	3
Math & CS	3	3	3	2	5	1	2	2	5	3
Engineering	15	8	4	7	3	5	6	6	6	4
Physical Sci	3	3	6	4	8	4	4	5	7	7
Social Sci	0	0	1	0	0	0	0	0	0	0
Did Not Report	1	1	0	0	1	1	0	0	0	0
Total	25	16	15	15	20	16	18	16	20	17

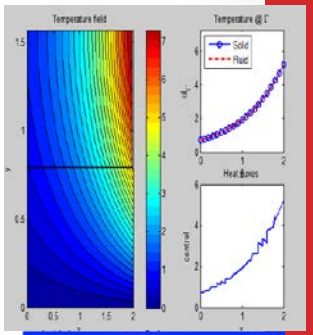
Judith Hill (CSGF Fellow, '99-'03)

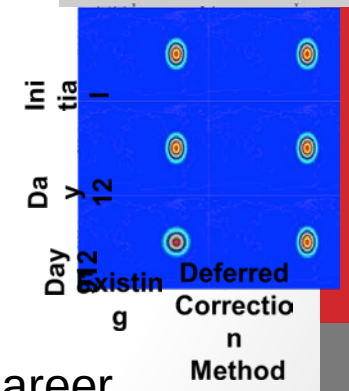
Computational Scientist and Liaison Task Lead
Oak Ridge Leadership Computing Facility
Oak Ridge National Laboratory

Research Interest – Numerical methods enabling the effective use of HPC resources

- Inverse problems; PDE-constrained optimization,
- Multiphysics coupling using optimal control
- Pseudo-spectral methods with multiwavelet bases
- Deferred correction methods for time integration
- Applications ranging across fluid dynamics, climate, nuclear physics, plant biology and chemistry





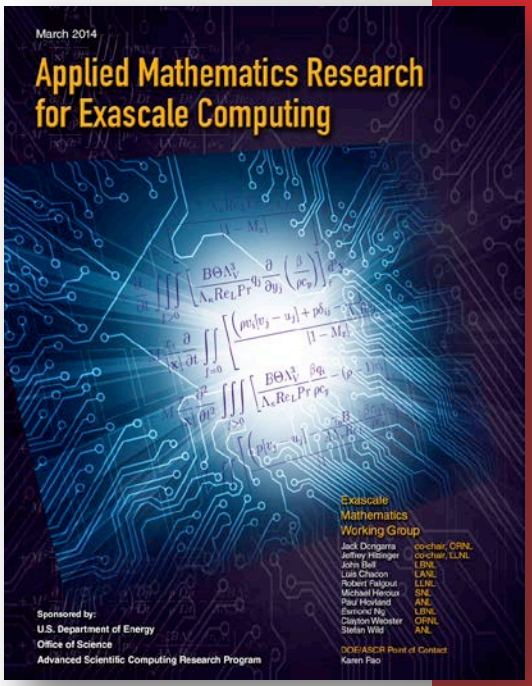


"The CSGF program of study encouraged me to pursue a career that was more computationally focused than I originally planned… These experiences directly influenced my choice of a career as a computational scientist in the DOE Laboratory system.."

CACDS

Jeff Hittinger(CSGF Fellow, 1996-2000)

Group Leader, Scientific Computing
Center for Applied Scientific Computing, LLNL

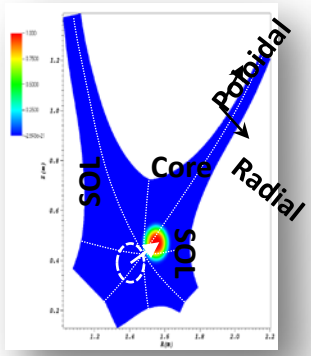


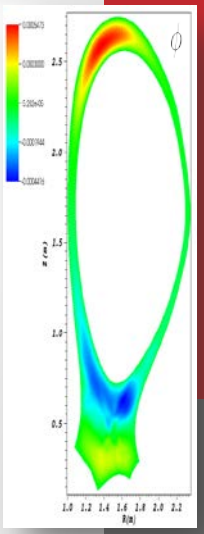
- Practicum: LANL, 1997
- PhD: U. Michigan, Aerospace Engineering and Scientific Computing
- Postdoc: LLNL, 2000-2001
- Research staff: LLNL, 2001-present

- Practicum:
  - Initiated his interest in working at a DOE lab
  - Contacts from practicum directly led to his postdoctoral position

- Program of Study:
  - Allowed him to pursue advanced training in applied mathematics and computer science
  - Breadth of training has made him a key contributor to several interdisciplinary projects

- Current and Past Leadership:
  - DOE ASCR Exascale Mathematics Working Group co-chair
  - DOE ASCR Applied Mathematics Point-of-Contact, LLNL
  - DOE OFES/ASCR Fusion Simulation Program planning team
  - Has served on several ASCAC subcommittees





CACDS

Existing Workforce Training

- Multiple reviews attest to success of CSGF program
  - Contributes to lab, national workforce in Comp. Sciences
- 2011 ASCAC review of the program states that:
  - "a large percentage of fellows spend a portion of their early career in the DOE laboratories and an even larger portion continue interaction with the DOE laboratories as they pursue their careers in academia and industry."
- 2012 longitudinal study of CSGF program:
  - 155 respondents: 28% in government, 38% in education, 34% in industry; 89% had CS&E-related employment

FINDING: The exemplary DOE CSGF program, deemed highly effective in multiple reviews, is uniquely structured and positioned to provide the future workforce with the interdisciplinary knowledge, motivation, and experiences necessary for contributing to the DOE mission.

Lab Training and Outreach

- Labs all put significant effort into attracting, training and retaining workers in Computing Sciences
- Engage graduate students and postdocs through programs such as summer internships, university subcontracts
  - Exposure to lab environment is key
  - Postdoctoral fellowships such as ANL's Wilkinson and ORNL's Householder
  - Programs where student is co-advised by lab staff
- Many staff come from these programs

DOE Workforce Development
Office of Science

ANL's ATPESC 2014: One of many lab initiatives in outreach and training

Role of Labs: Workforce Retention

- Labs face many challenges in maintaining workforce
- Need to re-examine career paths to offer competitive choices, provide a more attractive workplace
  - Current funding model makes it hard for young staff to establish themselves
  - Consider how to give employees opportunities to grow professionally e.g. through new opportunities
  - Provide resources to address work/life balance
  - Adapt to shorter-term commitment that is becoming common
  - Facilitate mid-career entry to labs
  - Engage in education

FINDING: The DOE laboratories have individually developed measures to help recruitment and retention, yet more can be done at the national level to amplify and extend the effectiveness of local programs.

Summary of Findings

- The multidisciplinary national labs face workforce recruiting and retention challenges in Computing Sciences
- Insufficient educational opportunities are available at academic institutions in areas of the Computing Sciences most relevant to the DOE mission.
- There is a growing demand for graduates in Computing Sciences that far exceeds the supply. A larger workforce gap and continued underrepresentation of minorities and females are expected.
- The exemplary DOE CSGF program, deemed highly effective in multiple reviews, is uniquely structured and positioned to provide the future workforce with the interdisciplinary knowledge, motivation, and experiences necessary for contributing to the DOE mission.
- The DOE laboratories have individually developed measures to help recruitment and retention, yet more can be done at the national level to amplify and extend the effectiveness of local programs.

Full Set of Recommendations (1)

- Preserve and increase investment in the DOE CSGF program to opportunities for more high-quality students, particularly students from underrepresented populations and demographics.
- Establish new fellowship programs, modeled after the CSGF program, for research opportunities in enabling technologies in the computing sciences, including computer science for HPC, large-scale data science, and computational mathematics.
- Expand support for local laboratory programs and encourage greater inter-laboratory sharing of information about locally successful programs and workforce related data.

Full Set of Recommendations (2)

- Establish a DOE-funded Computing Leadership graduate curriculum advisory group to spearhead participation in efforts within ACM, CRA and NSF to develop and annually publish competencies of DOE need at the graduate and undergraduate level.
- Working with ACM SIGHPC, NSF and other organizations, provide a rich repository of DOE mission-oriented learning materials and engagement opportunities to attract and guide individuals towards careers in areas of DOE need.
- Working with other agencies and organizations, establish certificate programs to address need for competency certification. Work with other agencies to fund implementation of curricular programs, particularly online programs, in the areas of DOE need.

Full Set of Recommendations (3)

- Improve attractiveness of DOE opportunities with continued relocation assistance, ongoing professional development in DOE strategic areas and position rotation, and establish a sabbatical program for DOE employees.
- Increase awareness of DOE opportunities by working with multiple universities to develop campus champions and increase support for DOE employees to visit campuses to promote opportunities within DOE.
- Working with other agencies, develop a strategic plan with programs and incentives to pro-actively recruit, mentor and sustain the involvement of significantly more women, minorities, people with disabilities, and other underrepresented populations through the completion of their PhD program and their active participation in CS&E careers.

Summarized Recommendations

- Leverage and strengthen the successful DOE CSGF program by doubling its funding and expanding its scope to include research in HPC-enabling sciences such as computational mathematics, computer science, and data analytics.
- Establish a DOE-funded computing leadership graduate curriculum advisory group for establishing graduate level curricular competencies specifically to fulfill DOE's Computing Sciences workforce needs.
- Develop a recruiting and retention program that increases DOE's visibility on university and college campuses and that provides relocation assistance, travel for recruiting, ongoing professional development, opportunities to take sabbaticals and other non-monetary career incentives.
- Build a diverse workforce that spans demographics and universal accessibility for a broader awareness of career opportunities within DOE. This includes inter-agency coordination, targeted opportunities, and strategic development plans to expand and build a workforce to include non-traditional, underrepresented, and military veterans.
- Expand support for local laboratory programs and encourage greater inter-laboratory sharing of information about locally successful programs as well as workforce related data.