HiPOP

EU/US ATLANTIS Programme

Cooperation in Higher Education and Vocational Training

December 2011

Grant Agreement Number: 2009 – 3195 – CPT EU – US POM

Deliverable 2
Comprehensive Policy Report

Version: 7.0
Authors: CyI – Klaus Schilling
CyI – Christos Nicolaou
CyI – Michael Frank
GRS – Boris Orth
BUW – Francesco Knechtli
JSC – Martin Sczimarowsky
NCSA – Scott Lathrop
NCSA – Danny Powell
Shodor – Bob Panoff

Date: 12/29/2011

Table of Contents

List of Acronyms

AICES	Aachen Institute for Advanced Study in Computational Engineering Science
BUW	Bergische Universität Wuppertal
CaSToRC	Computation-based Science and Technology Research Center
CDS&E	Computational and Data-Enabled Science and Engineering
CI	Cyber Infrastructure
CPU	Central Processing Unit
COMPASS	Computer Laboratory for Parallel Algorithms and Stochastic Simulations
CS	Computer Science
CS&E	Computational Science and Engineering
CSERD	Computational Science Education Reference Desk
CSiS	Computer Simulation in Science
CSS	Cascading Style Sheets
CST	Computational Science and Technology
CUDA	Compute Unified Device Architecture
DFG	Deutsche Forschungsgemeinschaft
EOT-PACI	Education, Outreach and Training Partnership for Advanced Computational Infrastructure
EU	European Union
FZJ	Forschungszentrum Jülich
GEMS	Girls Engaged in Math and Science
GIMP	GNU Image Manipulation Program
GIS	Geographic Information System
GPU	Graphics Processing Unit
GRS	German Research School for Simulation Sciences
HiPOP	High Performance Computing Opportunities
HPC	High Performance Computing
HTML	HyperText Markup Language
ICLCS	Institute for Chemistry Literacy Through Computational Science
JSC	Julich Supercomputing Centre
ITEST	Information Technology Experiences for Students and Teachers
IZII	Interdisciplinary Center for Applied Informatics and Scientific
MPI	Message Passing Interface
NCSA	National Center for Supercomputing Applications
NCSI	National Computational Science Institute
NEH	National Endowment for the Humanities
NSDL	National Science Digital Library
NSF	National Science Foundation
PRACE	Partnership for Advanced Computing in Europe
R&D	Research and Development
RTP	Research Triangle Park
RWTH	Rheinisch-Westfälische Technische Hochschule

Shodor	Shodor Education Foundation, Inc.
SIR	Susceptible, Infected, Recovered
STEM	Science, Technology, Engineering, and Mathematics
SUCCEED	Simulate Understanding of Computational science through Collaboration, Exploration, Experiment, and Discovery
UIUC	University of Illinois at Urbana Champaign
VSCSE	Virtual School of Computational Science and Engineering
VVA	Verification Validation and Accreditation
VVR	Verification, Validation, and Accreditation
WCU	Western California University

0.0 Executive Summary

Computational Science and Engineering (CS&E) have emerged over the past decades as a highly efficient methodology to understand and master complex systems – as they occur in advanced R&D - through computer simulation. Yet, at their very frontier, namely at the level of High Performance Computing (HPC), the critical need of R&D is severely impeded by a shortage in the available workforce. The HiPOP consortium therefore analyzed current local promotional practices of worldwide institutions to overcome this serious workforce shortage through recruiting and building the next generation of scientists, engineers, scholarly researchers and practitioners.

To effect large-scale alleviation in the above shortage, current and future educators must be well prepared to immerse their middle and high school and undergraduate students in CS&E for the long-term advancement of science, technology, engineering and mathematics (STEM). In particular, there is an urgent need for creating and providing suitable materials for these educators to better illustrate/demonstrate the potential of models, simulations, and visualizations in their classrooms. In particular, simulation experiments at school enhance inquiry-based learning and will significantly motivate more students to pursue STEM education and careers.

We need to both achieve and sustain large-scale preparation of students from middle school through their graduate education. We can accomplish this by adopting a vertical approach to education that provides a systematic learning continuum through collaborations among K-12 schools, colleges, CS&E experts, business and industry.

As a result of the HiPOP survey among professionals, we recognize that there is a broad enough base of organizations and experts willing to participate in such large-scale endeavors. Collaborations such as these can be directly enhanced through better tools to enable and enhance these capabilities. Furthermore, these types of tools need to be incorporated into the educational framework to support K-lifelong learning systems.

As an important step towards that goal, we present in the following a comprehensive policy report for better educational outreach by use of the Internet, that would both:

- create synergies among present outreach activities as practiced within the CS&E/HPC communities on both sides of the Atlantic and
- achieve higher and wider scalability of present activities worldwide to mobilize human resources for CS&E.

1.0 Introduction

HiPOP (High Performance Computing Opportunities) is a project funded by the EU-US Atlantis Programme, whose goal is to promote high performance computing (HPC) in computational sciences to the non-HPC community in order to attract young talent to this scientific domain. HPC is the high end of computational science and engineering (CS&E) and so young people should first be directed to the latter, from where they will be able to advance to HPC at a later stage.

The project is coordinated by the Cyprus Institute, Nicosia, Cyprus (http://www.cyi.ac.cy) with the following partner institutions:

- German Research School for Simulation Sciences (http://www.grs-sim.de/)
- Bergische Universität Wuppertal (http://www.uni-wuppertal.de/)
- Forschungszentrum Jülich – Jülich SuperComputing Center (http://www.fz-juelich.de/jsc/)
- National Center for Supercomputing Applications at the University of Illinois at Urbana – Champaign (http://www.ncsa.illinois.edu/)
- Shodor Education Foundation, Inc (http://www.shodor.org/)

This document constitutes the second deliverable of the project, which according to the proposal is a "Comprehensive Policy Report". Its purpose is to reconcile the partner policy reports, compiled during the first phase of the project, with information gathered from the HiPOP online community survey and from the "Berlin User Meeting" among others. Specifically, the information in this report comes from several sources including:

- A survey conducted as part of the HiPOP project
- The Partner Policy reports
- Discussions among the partner institutions
- The "Berlin User Meeting" organized by HiPOP
- Information and statistics from other relevant sources and surveys.

In order to prepare this policy successfully, the following steps were taken:

- The motivation for the HiPOP project as well as the sources of the related problem were identified. The best practices we selected for the policy presented in the current document have been chosen so as to directly address the problem.
- Efforts with goals similar to the ones of HiPOP have been identified. For previous efforts, advice has been taken as to what worked for them and

what did not. Current efforts have been recorded and ways of working to complement each other's work are taken into account.

- Finally, based on the previous items, a step by step strategy was specified that clearly shows how to reach the final goals of HiPOP and specifically how to reach, inform, excite and eventually attract young talent to the CS&E field.

The rest of this report expands on the points above.

The overall goal of the HiPOP project is to better attract the younger generation into CS&E and HPC. To achieve this goal, HiPOP aimed to first define a policy to address the stated problem and then engage in activities leading to the partial implementation of the policy defined. Recognizing that the problem is common to the EU and the US the project consortium combines partners from both sides of the Atlantic, thus enhancing transatlantic communication between EU and US HPC communities. Specifically, the objectives of HiPOP are to:

- Identify local best practices and possible solutions to the lack of response by students to newly developed Master and PhD Programs in CS&E, and
- Recommend and implement strategies for the deployment of WEB based tools for dissemination of information on HPC/CS&E.

The final product of the HiPOP project is a prototype of an HPC/CS&E WEB portal aims of which are to:

- raise public awareness and engagement among youth worldwide,
- provide access to high quality materials and mentoring, and
- stimulate creativity via global collaborations among students, educators and practitioners.

2.0 Overview of the Problem

Computational science is a research methodology that uses computer-based models and simulations to solve problems both in pure and applied sciences as well as in engineering. For many years now, the use of computers in science and engineering has assisted greatly in transformational advances that are benefitting society. Being able to model and simulate complex systems in a *virtual* experiment, i.e. by computer simulation, is in many cases more effective than actual experimentation (and in cases which are too expensive, too dangerous, or of an unrealistic scale to replicate, simulation is the only available option, like in astrophysics). Here are some few examples for illustration:

- Complex systems like Many Body Quantum Physics/Chemistry
- Quantum Engineering in Materials Science
- Collective phenomena of thermodynamic systems, like phase transitions
- Drug design
- Climate research
- Astrophysical systems
- Aircraft optimization
- Nuclear systems designs

In many applications at the frontier of science and technology, the computational needs are extreme and hence progress is *capacity* bounded. As capacity expands, a couple of options open – the complexity of the simulation can be enhanced (to better control systematic errors), or more parameter sweeps (ensembles) can be run (to increase accuracy). The quest for more realistic simulations motivates large scale influx of funding for HPC systems world-wide, with an active effort to provide petascale production class systems, while research is underway to develop exascale class systems. In this area of high end computing, computational power is achieved nowadays by the use of massively parallel computer systems, with many compute nodes working together to simulate the system at hand in due time, along with the development of new algorithms, tools and applications. The science of adjusting a program in such a way as to best utilize multiple resources (clusters, CPUs, GPUs, networking, data storage, large scale data analysis, etc.) is called high performance computing and constitutes the high end of computational science and engineering. It also extends into non-traditional fields of research including humanities, arts, and social sciences.

The *main concern of the HiPOP* project is that even though computational sciences and high performance computing offer proven windows of opportunities, both for research and for industrial purposes, the number of students involved in such topics is not keeping pace with the needs of computation-based science and technology. This is evidenced by the fact that an entire sequence of sessions of the International Supercomputing 2010 and 2011 Conferences were devoted to the Special Topic :

“The HPC workforce Development” (see for example http://sc10.supercomputing.org/?pg=specialtopicHPCWD.html)

One of the obvious causes of the recruitment problem is that young people are not sufficiently exposed to and informed about computational sciences. Most students just do not know about the broad scope of CS&E and its value to society. They tend to identify it with computer science rather than viewing it as an opportunity to make sophisticated use of computers to assist them in their application domains.

It goes without saying that this human resource shortage in computational science is part of a more general problem, namely the serious shortage of young talent entering the STEM disciplines. A 2009 report from The Russell Group of Universities in the UK for example states that during the period of 1989-2004, the number of students taking math fell by 40%. In 2008, the percentage of students taking STEM subjects increased slightly but the numbers are still below their previous levels. The same report indicates that one in four employers had experienced difficulties in hiring STEM graduates.

Last but not least, students—especially the ones studying natural sciences and mathematics—become at some point suddenly involved in the world of programming, which with no previous background in computer science or without adequate motivation seems very strange, complex and frightening for students. This definitely discourages them from becoming immersed in research projects whose progress is largely dependent on sophisticated algorithm and software research to attain higher efficiencies on scientific simulations on ever advancing hardware systems spanning computing systems, networking, storage systems, visualization systems, and desktop HPC tools.

Motivated by the above findings and conclusions, the HiPOP team investigated and documented the best strategies for attracting young talent to the CS&E field, and, designing and implementing ways for exploiting WEB-based methods to expand the potential to recruit and sustain significantly larger numbers of students.

3.0 Review of Previous and Current Attempts

To accomplish this project, the HiPOP team decided to start by understanding and identifying current activities that are used to promote computational sciences and high performance computing as well as methods used to publicize these activities. The institutions of the international HiPOP consortium provided much of the primary information, by preparing individual policy reports during the first deliverable of the HiPOP project. To further broaden its basis of the understanding of the challenges and opportunities, the consortium decided to conduct a survey among a diverse range of institutions from the EU and the US active in CS&E.

This section starts by briefly quoting the essential recommendations from our previous partner policy reports and then expands on the main findings and messages of the HiPOP survey. Both these information sources serve in later sections of this report as the basis for a comprehensive policy report.

3.1. Partner Policy Reports

The consortial partners represent an extensive dissemination network for SC&E/HPC, ranging from outreach to high school (to alleviate the STEM problem) through tertiary education and research school levels to HPC centers. From their outreach policy reports, there is clear evidence that:

- we witness a serious reduction in student numbers in computer science and CS&E in Western countries;
- we see the need for an international CS&E community formation to enhance CS&E visibility among the youth;
- there is plenty of room to improve the standing of CS&E among the youth of under-represented communities;
- we see plenty of opportunities to attract more talented youth from pre-industrial countries;
- there is a definite need for an effective way to disseminate CS&E to the youth by demonstrating its educational merits within the training of next-generation teachers;
- Internet based methods for promoting CS&E promise high outreach in complement to proven face-to-face methods

3.2. The HiPOP Survey

Based on the experience of our HiPOP Partner Policy Reports (see also Appendix B) the HiPOP survey addressed itself to individuals organizing activities (which are mostly academics and personnel of HPC centers) for promoting computational sciences and HPC. Its overall goal was to gather information that will help us in defining best-practices relevant to attracting students to CS&E and in designing an online educational portal, a cyber-platform, with appropriate functionality and content. The questions of the survey, which were carefully devised by the consortium, aimed to retrieve information on:

- the person filling the survey (to initiate worldwide community formation)
- the activity carried out by that person (to gain an activity profile)
- the participants of the activity (to possibly identify additional target audiences for educational WEB activities)
- personal opinions on how to arrange a successful activity (to optimize WEB outreach strategies)
- what is available and what is missing from the web, with respect to educating and informing people in HPC (to assess current practice of the WEB for HPC education)

To foster brainstorming about the future cyber-platform, we decided to incite novel ideas/suggestions by avoiding the exclusive use of multiple choice style questions. Instead, for the more important ones, we allowed the users to expand their answers as they deemed necessary. This narrative mode, as expected, made the analysis of the survey considerably harder. A full listing of the survey questions can be found in Appendix A. The survey was distributed online, via the HiPOP project website and run between 18/05/10 – 7/06/10. The survey was advertised among the computational science educational community and HPC centers.

With feedback from more than *100 completed surveys* our analysis took place in three stages:

1. The first stage involved analyzing the multiple-choice type questions, i.e. questions in the form of drop down menus, check boxes and radio buttons. These are restricted to options that we specified beforehand and hence, analysing them was straightforward. The data extracted were then exported into a spreadsheet and presented visually in the form of

bar charts (one bar char for each question where each bar represents an option for that question).

2. The second stage involved analyzing the answers to the more difficult questions; i.e. those in which the users were allowed to freely enter the information they wanted inside a text box. We found that most of the answers for each question could be manually grouped into a small number of categories.
3. Upon stage 2 the final analysis followed the usual pattern.

The detailed results are presented in Appendix A.

Let us quote here our main findings and *messages in point-wise mode:*

A. The majority of respondents were faculty members in educational institutions. *This provides us with a reasonably large base of experts to cooperate with for fostering HPC.*

B. The survey was answered mostly by males (almost twice as many as females), reflecting male dominance in the field: *there is a critical need to attract female students into the field.*

C. Most reported activities promoting HPC are organised by people who are older than 45 years old. It would seem that younger academics are less likely to organise such activities or participate in surveys addressing the challenges in motivating HPC. *We infer that it would be worthwhile to create more interaction with young researchers, at postdoc and PhD student levels, for bringing up new ideas for a better appeal of CS&E to the youth. To this end we would like to see HPC education earn more academic credentials and recognition.*

D. The vast majority of respondents are Caucasian, which is interesting since over half of the institutions that answered the survey were US institutions. *This illustrates once again the challenge to broaden participation in CS&E. On the EU side, we see a point in bringing the message to young talents from the Southern and Eastern neighbourhood areas, which at present are HPC diaspora areas. On the US side, we need to see more participation by women and minorities.*

E. Most of the activities promoting HPC are organized by universities rather than by other institutions. This is plausible as University students have courses and research projects that utilize CS&E methods. But it also reflects the fact, that CS&E is not an integral component of K-12 education and that an inadequate number of in-service and pre-service teachers are being prepared to incorporate CS&E methods and resources into math and science courses to prepare the next generation of computational scientists and engineers.

F. Promotional activities are mostly targeting college students, with greatest emphasis at research level institutions, secondarily at primarily undergraduate institutions, and with least effort among two-year colleges and universities. This is another facet of the shortcoming mentioned above under E and reflects the motivation to recruit from existing college students - that already have experience in their science, and can therefore readily appreciate the benefits arising through the use of computing and HPC in their domain of study - rather than *to head for a broader baseline and demonstrate/illustrate the potential of computer simulations during the very basic teaching of science.* To illustrate the point: with CS&E methods the notorious problems of teaching classical dynamics in analytical form can be circumvented: the CS&E approach amounts to use of operational, discrete mathematics and is much easier to grasp for the high school student in his learning process, both in Math and Physics.

G. Most local activities as reported in the survey are not sustainably funded. *This hints to the importance of using WEB technology for disseminating best practices and material – and to foster coherent actions by those means. But one should keep in mind that the issue of sustainability remains for Internet resources (such as a HPC-community educational portal) as well, though – seen globally - at a much lower cost level than the overall costs of individual, local dissemination activities.*

H. Most activities covered are of an educational nature, are repeated frequently rather than being one-offs, and continue throughout a whole semester – suggesting that they are probably part of established educational programs. We can state clearly that many universities worldwide created such programs in CS&E over the past 1 – 2 decades, with many of them around close research groups with CS&E applications*. Nevertheless quite a few of these educational initiatives still function in non-established mode, i.e. with only initial funding support.*

I. We find both, regular activities within formal courses with more than 100 participants, and informal education projects which tend to be occasional and to have fewer than 30 participants.

J. The vast majority of activities are conducted in English. *This points to the issue whether one should plan for promotional actions in other wide-spread languages. We have seen some efforts to translate materials to other languages, but at a slow and inadequate pace to meet international needs.*

K. In most cases, participation in activities is live, which limits the outreach in space and time. This again calls for the need of better using WEB infrastructure, both to exploit synergies and to counteract a digital divide: breaking down the barriers to allow everyone full and equal access via the web to learning opportunities. Many of our traditional instructors and workshop leaders need to learn how to work in an on-line (synchronous and asynchronous) mode. Efforts to migrate events from live events to on-line live events have been met with opposition or have been poorly delivered by people who have not been well prepared in conducting such efforts.

L. In most cases the majority of participants were males between the ages of 18 and 24, consistent with the overall make-up of the CS&E student population. *This corroborates that there is plenty of room to improve the outreach towards high schools, women, minorities, and people with disabilities.*

M. The majority of HPC educational activities were promoted via email, with WEB sites and portals playing no significant role. On the other hand, the Shodor WEB portal that addresses the K12 world reports monthly access rates of more than 3 million on their CSERD digital library. Hence it appears that the *opportunity in exploiting the Internet for the sake of HPC education has been missed so far.*

N. The most common measure for assessing the success of activities was through student evaluation forms. In some cases student grades were also considered. Most projects are of a short-term nature and lack longitudinal studies of impact.

O. Frequently mentioned success factors included "hands-on" activities, quality of teaching staff and students.

P. Few respondents make significant use of existing CS&E and social networking sites. There is only marginal such HPC social networking activity at Shodor and HPCU with Facebook. It is expected that an appropriate *social networking feature on an HPC-education portal would add an important dimension to the promotional outreach, to the benefit of a better understanding and appreciation of HPC among the youth.*

Q. Very few respondents answered a question asking which WEB functionality they believe would enhance their work and activities. This suggests that most have not really thought about this aspect of promoting their activities. *The field is still in its infancy.*

The overall conclusion drawn from the survey analysis in conjunction with the consortium partners experiences is that it appears both necessary and timely to implement WEB based methods in order to attain more coherence between worldwide efforts for improving the CS&E/HPC workforce situation. Given the scarcity of sustained funding, this can best be done by using all possible synergies, in particular by concrete cooperation between EU and US, and enhancing existing programs like the one of the *Community for Science Education in Europe* SCIENTIX with outreach efforts of the CS&E research communities worldwide.

4.0. Comprehensive Policy - Discussion

The aim here is to develop a set of general guidelines in the form of a comprehensive policy for educational outreach in CS&E.

The latter should be seen as a strategy space, i.e. a collection of all the available options for inciting a next generation of scientists and engineers in favour of CS&E. The best suited ones of those options should be considered in designing efficient international projects to promote education in computational science and HPC (and disseminating them through a community driven WEB portal). The actual choice for the present HiPOP project will be based on best practice, as emanating from the partner policy reports and the survey results.

In view of the available material we have decided to organise the policy guidelines along the following main categories:

1. Classification and selection of target audiences
2. Kinds of activities to be considered
    a. Educational activities
    b. Promotional activities
3. Methodologies for organizing promotional activities
4. Ways of evaluating the success of activities.

4.1. Classification of Target Audience

Target audiences can be classified according to a number of characteristics. The most relevant of those are by educational level/age group, occupation/ specialization and other characteristics such as sex, nationality, ethnicity.

Many of the CS&E and HPC education, outreach and training activities have been focused on targeting the higher education community (college, university). Research shows however, that many young people lose interest in math and science during their K-12 years (particularly during their middle school and high school years). A number of projects have statistical evidence that the use of computing to support inquiry-based math and science projects enhance the likelihood of sustaining student interest in learning and pursuing math and science studies at all education levels (see e.g. the article of John P. Kubicek in the Canadian Journal of Learning and Technology, http://www.cjlt.ca/index.php/cjlt/article/view/149/142)

However, at present there are not enough educators prepared to introduce and sustain these kinds of activities in their classrooms across the learning continuum from K-12 through graduate level education. All too often, K-12 students are introduced to exciting and motivational inquiry-based learning opportunities at one grade level, only to find that at the next grade level, the instructors revert to more traditional teaching styles that do not sustain inquiry-based learning. As a result, the pipeline is inadequate to meet the workforce needs.

Projects that engage educators and students spanning K-12 and higher education provide a more effective learning continuum to support and sustain student motivation and interest in math and science, like the MIT+K12 project (which targets actually to get MIT students involved as well, see http://k12videos.mit.edu/). This kind of approach provides a more effective feedback loop from K-12 through higher education. This includes the training of future teachers (pre-service education) during their undergraduate education, who are then well prepared to reinforce computer-based inquiry learning when they join the ranks of K-12 teacher.

As part of the K-12 and higher education continuum, it is important to provide opportunities for pre-service and in-service teachers to work together – the pre-service teachers can bring new ideas into the classrooms of in-service teachers, and the in-service teachers can better prepare the pre-service teachers for their classroom experience.

It is also important to try to attract individuals in disciplines that are not traditional users of HPC; an example is NCSA's attempts to engage the interest of people in humanities, arts and social sciences. The more usual approach of concentrating on traditional user communities such as physicists, chemists and biologists creates an artificial barrier to expansion of HPC into non-traditional CS&E communities.

Another important consideration is to increase the size and diversity of the computational science/HPC community by targeting under-represented groups in the field, such as females, minority ethnic groups and young talents from geographical HPC diaspora countries. A number of projects are currently addressing this requirement (like PRACE and LinkSCEEM in EC and XSEDE in the US). Yet it is obvious that further efforts in the HPC educational direction are indispensable. In the words of Dr. Richard Tapia, Rice University:

> "No first world nation can maintain the health of its economy or society when such a large part of its population remains outside all scientific and technological endeavors."

Finally, the opportunities presented by CS&E/HPC result in a change of the workforce needs. Hence, it might be advantageous to retrain professionals to adapt to these new needs.

4.2a. Kinds of Educational Activities

The educational activities we have encountered fall into the following groups:

1. Courses formally included in science, technology, engineering and mathematics (STEM) curriculum at bachelor's level with at least the basics of computational science and engineering
2. Computational Science/HPC related master's and PhD certificate and degree programs

3. Short courses (e.g. weekend and summer schools) for students and/or professionals
4. Lectures, presentations and seminars offered on campus, at conferences, and other venues
5. E-Learning initiatives (e.g. online course material and tutorials for self-paced learning).

Reportedly, in the case of courses in bachelor's level curricula, it is important to find ways of making these appear more relevant and approachable to the students. The more students are exposed to such courses, the more likely they are to follow up with related Master's and PhD courses, or to be aware of and use their CS&E/HPC knowledge in their chosen careers. Master's and PhD level programs in CS/HPC do exist at a number of educational institutions but there is a need to create more of them, and also to ensure that their relevance to non-Computer Science graduates is more explicitly explained. Short courses can play an especially important role in bringing CS/HPC knowledge to people who were not exposed to this material during their formal education. And finally, individual lectures, presentations and seminars are believed to arouse interest and hopefully persuade people to pursue further information on the topics involved.

A conclusion that can easily be reached is that the CS&E community needs to foster more formal and informal educational activities at the secondary (high school) levels, while exploring opportunities to engage and prepare students at younger ages. Currently though, these levels seem to be more suitable targets for promotional rather than explicitly educational activities; educational activities will rather be focused on introducing more general science, computing and mathematics concepts with the aim of tempting students to enroll in related degrees.

4.2b. Kinds of Promotional Activities

The most common kinds of promotional activities reported are:

1. Presentations and demonstrations at open days, student job fairs, science fairs, and other special occasions at universities and other institutions
2. Competitions
3. Inter-institution collaborations and student exchange programs
4. Institutional WEB sites and portals
5. Electronic mailings
6. Poster and brochure mailings, newsletters
7. Visits to schools and other institutions
8. Word of mouth.

All of these approaches have been tried by partner and other institutions, but the evidence as to their effectiveness is limited and largely anecdotal. This holds in particular for outreach to school actions which present a serious large-scale problem, where we should seek for synergies with research groups with past experience in this field. In designing and implementing activities of this kind in the future (in form of local pilot projects and WEB based) it is therefore extremely important to incorporate ways of measuring their effectiveness as objectively as possible *and to distribute the results of such assessments widely.* Cost also needs to be taken into consideration: in the long term one would expect electronic means such as WEB sites and portals to be more cost-effective than people-intensive means such as word of mouth communication, but the challenge we have to face is the "build it and they will come" mentality as there are thousands of WEB sites working to attract students.

4.3. Methodologies for Organizing Activities

As mentioned, the kinds of activities discussed above have been tried by partner and other institutions but in general *there are no guidelines on how to set up, execute and evaluate a particular kind of activity*. Thus anyone wishing to establish some kind of activity must effectively start from scratch and re-invent the wheel.

In the case of formal educational programs and courses the generic principles applicable to all educational activities will apply regardless of subject matter; these are fairly well understood. For more informal activities, and especially for promotional activities, a set of guidelines for each kind of activity would be extremely useful. For instance, it is highly recommended to make systematic use of young tutors in these informal activities. To this end, organizations setting up new activities should enumerate at the outset the principles that will be applied, and should attempt to evaluate those principles – or at least identify a set of lessons learned – by including the use of social networking among the youth. A central repository of such knowledge should be created and maintained with the objective of ultimately creating a useful and relevant body of knowledge and best practices.

4.4. Evaluation of Activities

Survey responses as well as individual partner reports indicate that objective evaluations of the cost and effectiveness of activities are rare; with the exception of formal educational programs and courses, where there is a history of such evaluation, such data is practically non-existent. Yet, without such data each activity starts from the same baseline; there is no possibility for systematic improvement.

As we discussed above in the case of methodology, this is largely a matter of ensuring that every new project sets out by specifying explicitly an assessment strategy to be followed throughout the project. Again the results should be collated into a central repository that can be made available to individuals and institutions planning new projects. *Funding agencies could play an important role here*, by insisting that each new proposal includes an assessment strategy and perhaps also by providing the means for collating and distributing the results.

5.0. Comprehensive Policy - Definition

Based on the above discussion, the HiPOP team finds that the use of WEB technology for education, outreach and training is still in its infancy. And yet, the WEB has clearly demonstrated tremendous potential for exciting, engaging and preparing a significantly larger and more diverse HPC workforce among current and future generations. The HPC educational portal, as being developed by the HiPOP partners, is both timely and uniquely positioned for directly contributing to and realizing this potential internationally.

The following are key policy recommendations for *exciting, engaging and empowering a significantly larger and more diverse HPC workforce.*

1. There is urgent need to further collect and disseminate lessons learned, best practices, and quality reviewed materials from prior and current education, outreach and training initiatives on a global scale. These resources will allow the international community to build upon what has been learned to further advance HPC knowledge and scholarly education and research.

2. Developing a larger and more diverse HPC workforce is a global issue that can best benefit from international collaboration. The issues are in general very similar across countries and across continents. HiPOP has demonstrated that there is a great deal to be learned from one another. Every effort should be made to further encourage and support international cooperation to significantly expand the workforce.

3. The greatest benefits can be achieved by working with a range of students spanning middle school through the graduate level. Research shows that students begin to lose interest in science and mathematics at the middle school level, and the pipeline continues to shrink in each subsequent year of their education.

4. To effect large-scale change, current and future teachers must be well prepared to immerse their middle and high school students in CS&E for the long-term advancement of science, technology, engineering and mathematics (STEM). There is an urgent need for creating and providing suitable materials for these teachers to better illustrate/demonstrate the potential of models, simulations, and visualizations for motivating significantly more students to pursue STEM education and careers.

5. We need to both achieve and sustain large-scale preparation of students from middle school through their graduate education. We can accomplish this by adopting a vertical approach to education that

provides a systematic learning continuum through collaborations among K-12 schools, colleges, CS&E experts and business and industry. Through the HiPOP survey among professionals, we recognize that there is a broad enough base of organizations and experts willing to participate in large-scale endeavors. Collaborations such as these can be directly enhanced through better tools to enable and enhance these capabilities. Furthermore, these types of tools need to be incorporated into the educational framework to support K-lifelong learning systems.

6. No nation can fully realize it's potential if a significant portion of its population is precluded from full participation in the use of advanced technologies including access to computational science and engineering tools and resources. Every effort must be made to engage under-represented communities including women, minorities, and people with disabilities. These under-represented communities also include fields of study that have not traditionally been involved in CS&E and HPC such as humanities and social sciences. There are numerous examples amongst these communities that are now realizing the benefits of CS&E and HPC to advance their education and research.

7. Every effort should be made to interact with young researchers, at graduate and postdoc levels, to breed new ideas for a better appeal of CS&E to the youth and to induce CS&E community formation. This will entail understanding what excites and engages them through multi-media interactive WEB technologies, which are already prevalent in their daily lives, and expected to catch their attention.

8. Every effort should be made to provide students with a sense of long-term career opportunities available to them if they pursue advanced CS&E and HPC studies and then careers.

9. The increased utilization of Social Networking opportunities will help to attract today's students and foster collaborations and innovations. These approaches need to be studied to understand and further improve upon these innovative collaborative environments for engaging today's youth.

10. Every effort should be made to address and support global communities. This should include support for multiple languages and diverse cultural perspectives.

Academia, business and industry, government and non-profit organizations must collaborate to pursue these recommendations. Alone, these organizations will effect incremental change. In collaboration, they can realize large-scale transformational change for long-term impact and benefit,

6.0. Outlook: Towards the Realization of the HiPOP Portal

The general policy statements discussed in this document and summarized in the specific policy measures stated above, will form the background for the next (design and demonstrator) phase of the HiPOP platform project. We shall ensure that each of the issues stated by the policy is explicitly addressed in the *design* of the HiPOP cyber-platform.

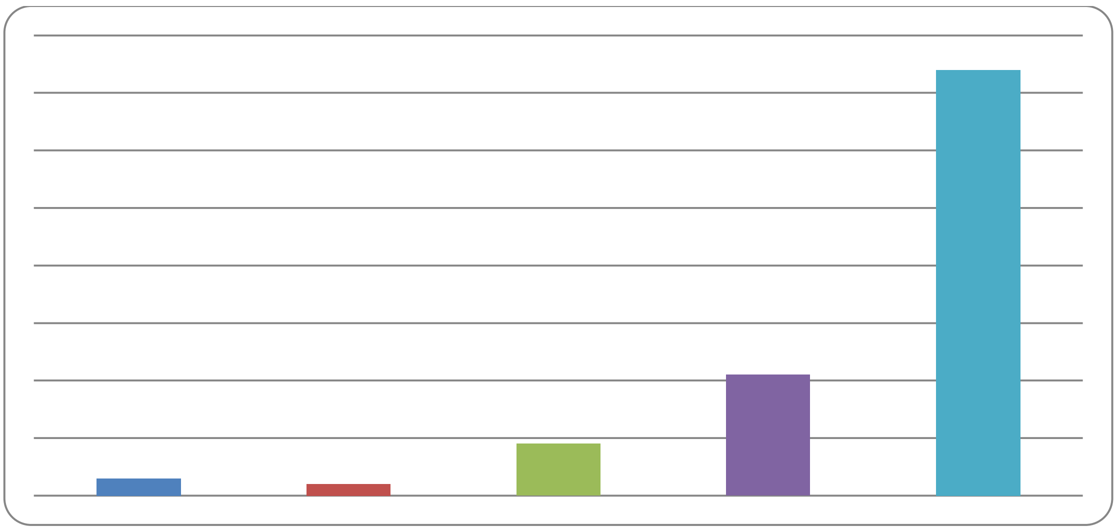
However, it is obvious, that the limited resources of the current project will not allow for the *full implementation* of the design produced; rather a core set of features is being identified and will be implemented following dialogue among the project partners and prospective users at various opportunities presented. This implementation strategy has been adopted in order to ensure that the tool developed will be in line with the objectives and budget of the project, the needs of the CS&E community and the identified best practices.

It is worth pointing out that the design specifications of HiPOP portal will foresee functionalities that will allow it to act as the much needed worldwide central repository of educational activities and guidelines for their best preparation with the objective of ultimately creating a useful and relevant body of knowledge and best practices. This feature, in combination with the comprehensive design of the platform will enable the phased release and completion of the tool through the continued cooperation and activities already planned by the project partners.

Appendix A – Survey Analysis: Detailed Results

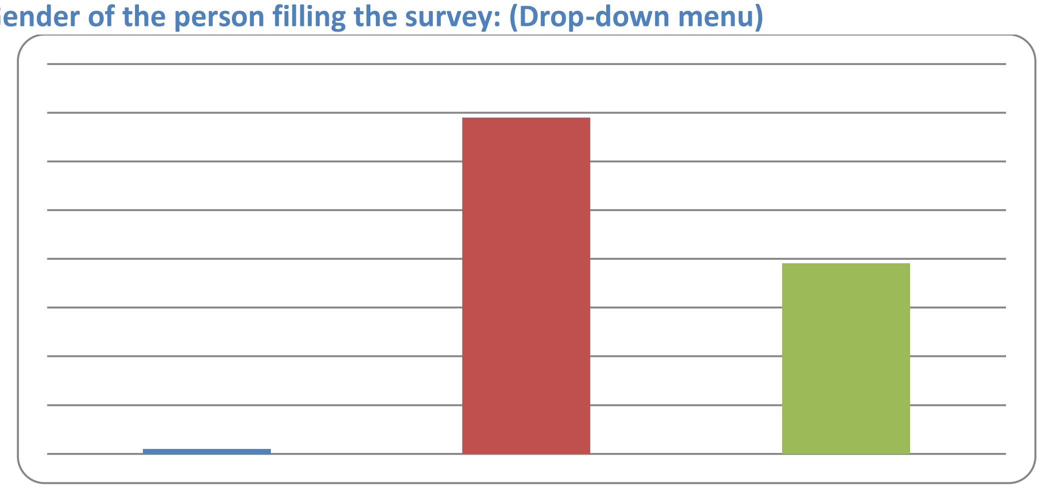
The entries in following distributions appear in the order of the drop down menus and check boxes of the HiPOP survey as published in the HiPOP WEBsite ,www.hipop.cyi.ac.cy

Age of person filling the survey: (Drop-down menu)



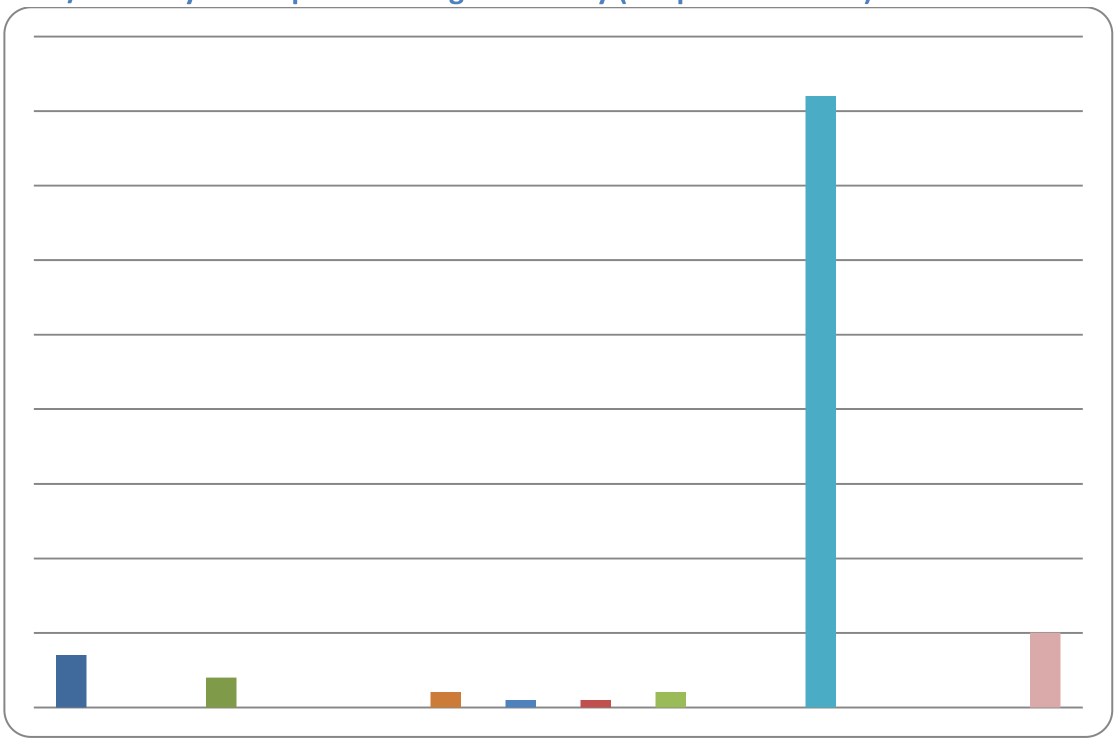
The above graph shows the distribution of the age of the person filling the survey. From this visual representation, we can see that the majority of activities for promoting HPC are organised by people which are older than 45 years old. Subsequently, we can infer that the younger the academic is, the less likely it is for him/her to organise such activity.

Gender of the person filling the survey: (Drop-down menu)



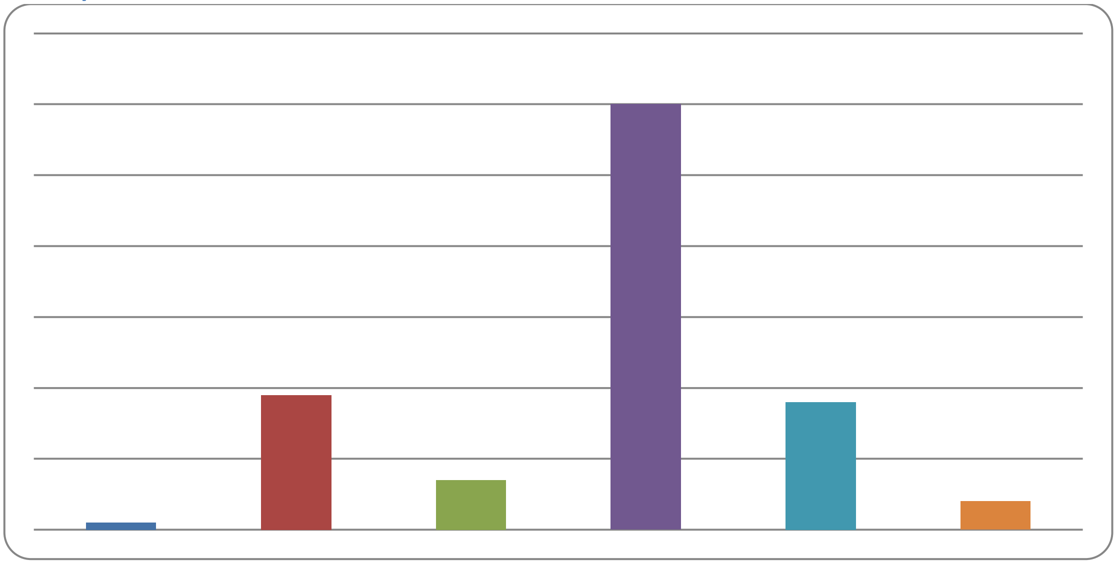
The above graph shows the distribution of the gender of the person filling in the survey. From it, we can see that the survey was answered mostly by males rather than females (almost twice as many males). This of course, is due to the fact that males make up a larger proportion of academics in the CS&E field.

Race/Ethnicity of the person filling the survey (Drop-down menu)



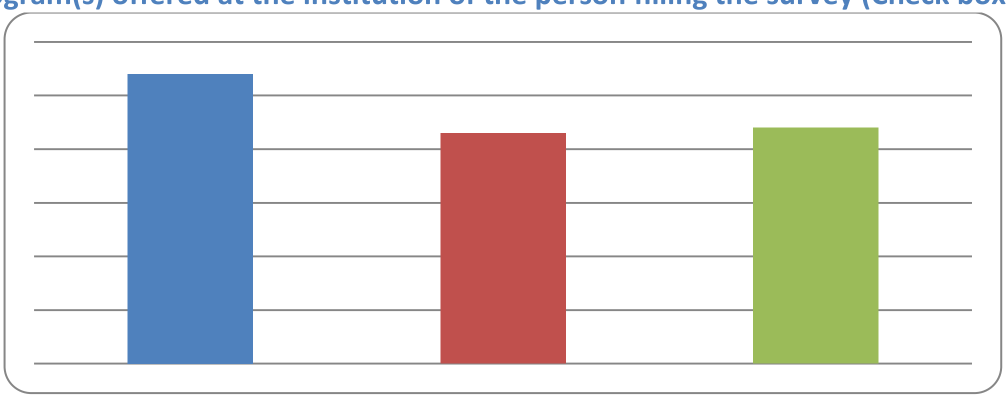
The above graph shows the ethnicity of the respondents of the survey. It is clear that the vast majority of them are European Caucasian, which is interesting since over half of the institutions that answered the survey were US institutions.

Position of the person filling the survey within his/her institute (Drop-down menu)



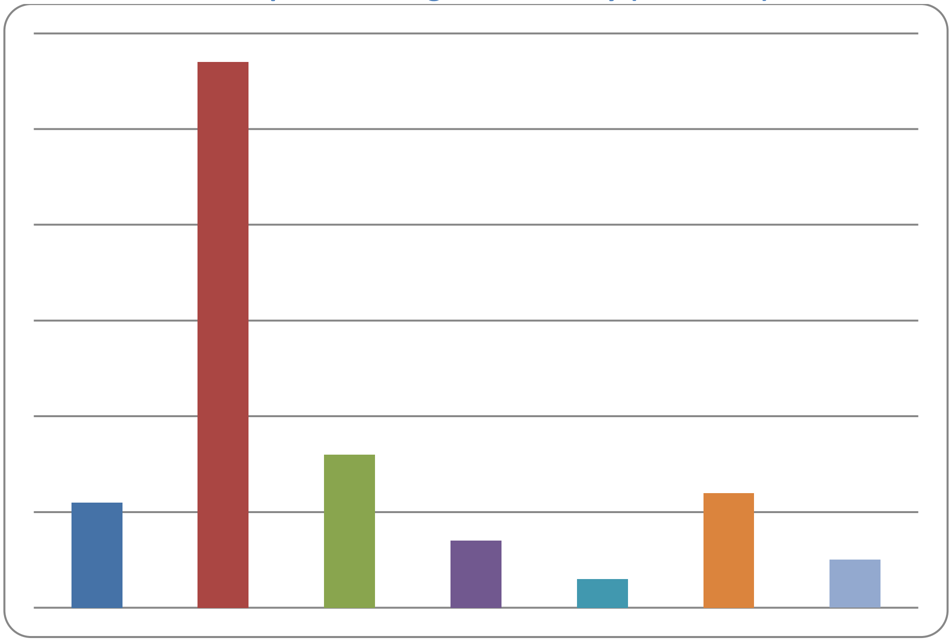
The question here is asking the position of the respondent of the survey within their institution. The results show clearly that the majority were faculty members.

Program(s) offered at the institution of the person filling the survey (Check boxes)



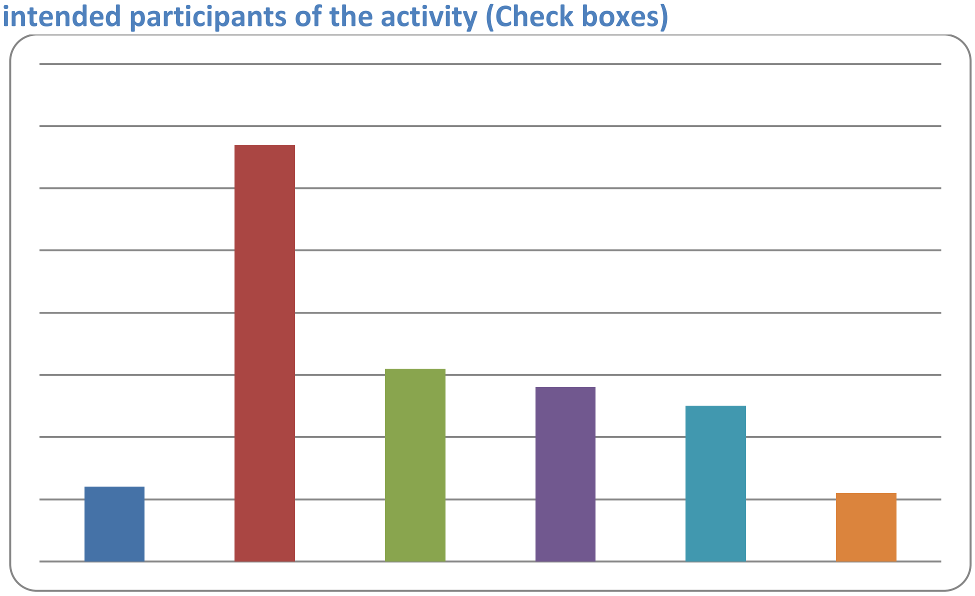
This question asks the respondent of the survey to state whether their institution offers a doctorate's degree, a master's degree or a bachelor's degree. Of course, an institute can have zero or more of the above options, so the total number of answers shown in the graph do not add up to the number of respondents of the survey.

Type of institution of the person filling in the survey (Free Text)



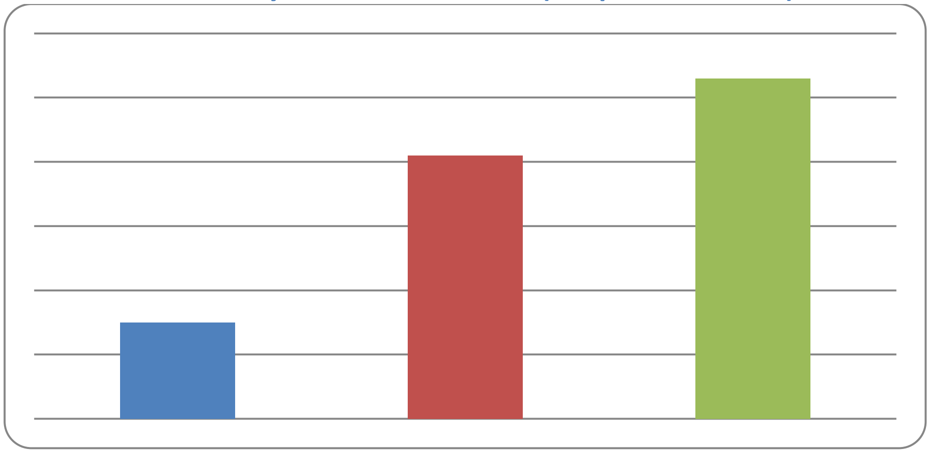
The above graph shows the type of institution of the respondent of the survey. The results show that most of the activities promoting HPC are organized by universities rather than other institutions. In some cases that might be true, as University students might appreciate such topics more than K-12 students. However, this could be due to sample bias: the survey was sent to many friends of the consortium partners, which are mostly universities.

The intended participants of the activity (Check boxes)



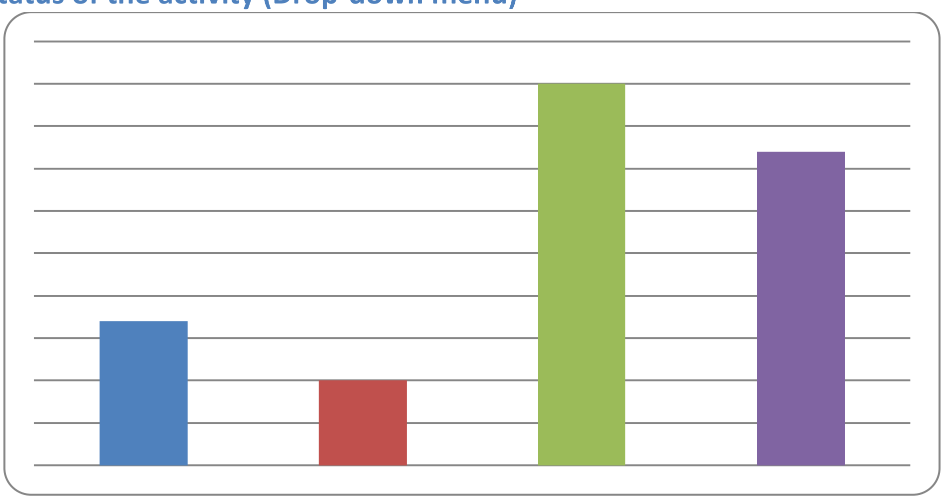
From the above graph, it is clear that these promotional activities are mostly targeting college students. This makes more sense, since college students have experience in their science, and can therefore realise the benefits arising through the use of computing and HPC in their domain of study.

Shows whether the activity was funded or not (Drop-down menu)



This question asks the respondents whether they received any financial help in organising their activity. The results show that such activities are not usually funded and that individuals preparing such activities are usually on their own.

The status of the activity (Drop-down menu)

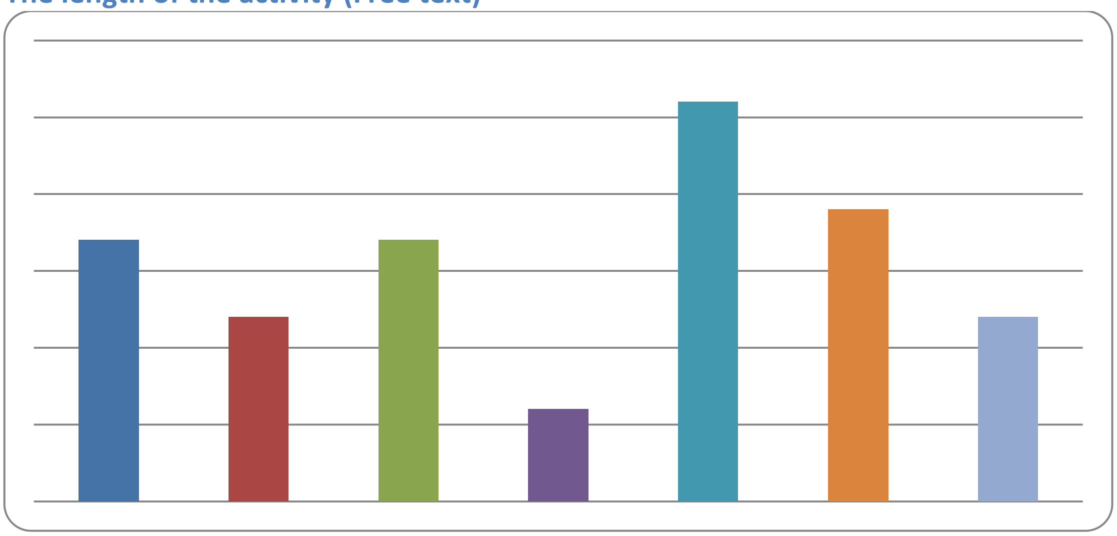


The bar chart above shows that during the survey period, most activities were still ongoing.

The frequency in which the activity occurred (Drop-down menu)

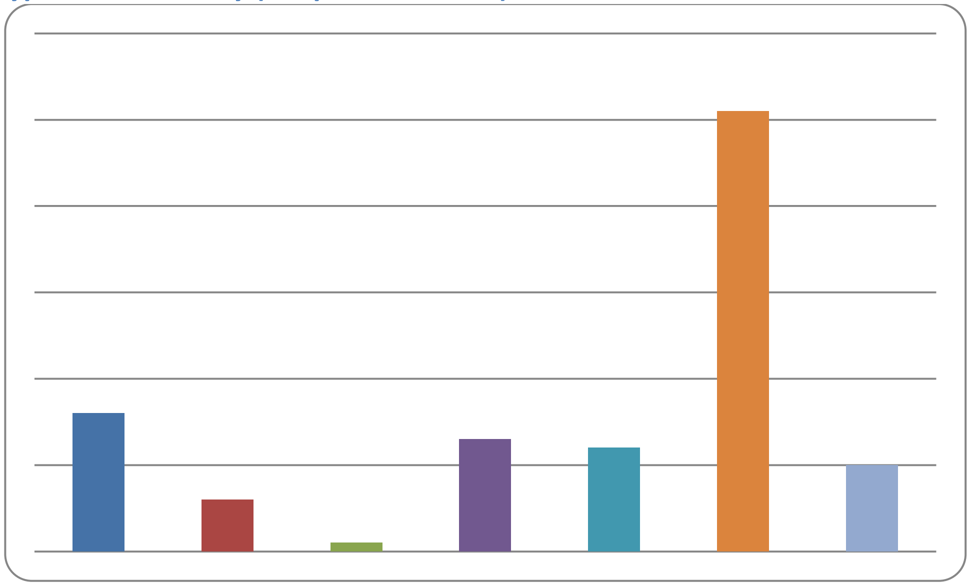
This question is asking the participants of the survey whether their activity was only organised once or if the same activity has been prepared multiple times. Obviously, institutions tend to arrange these activities frequently.

The length of the activity (Free text)



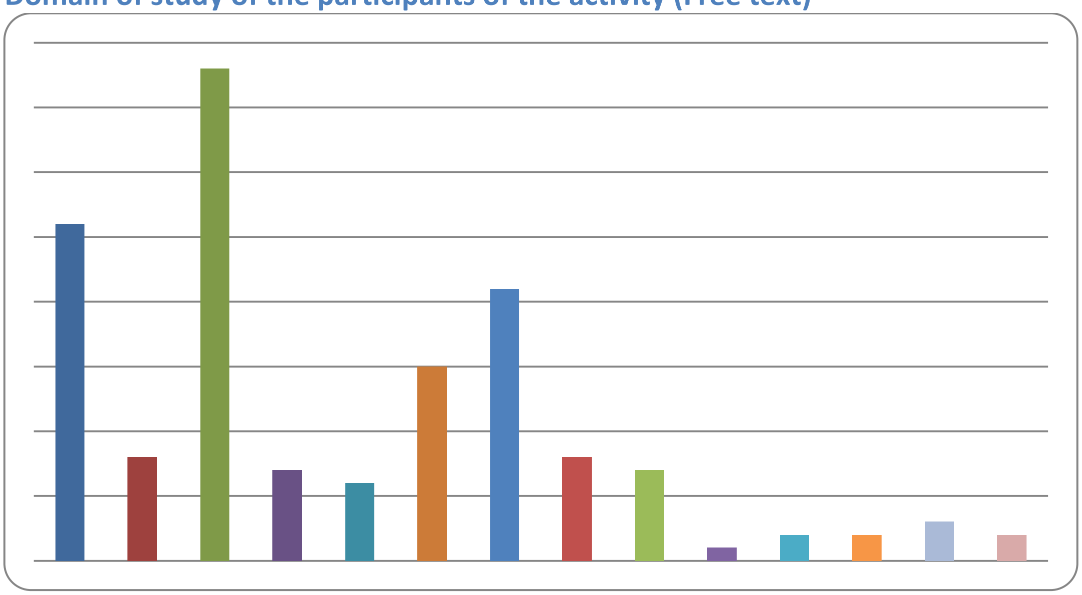
This question asks the respondents of the survey to state the duration of each instance of their activity. It seems that most activities continue throughout a whole semester, which makes us think that it is probably part of an institutional course.

The type of the activity (Drop-down menu)



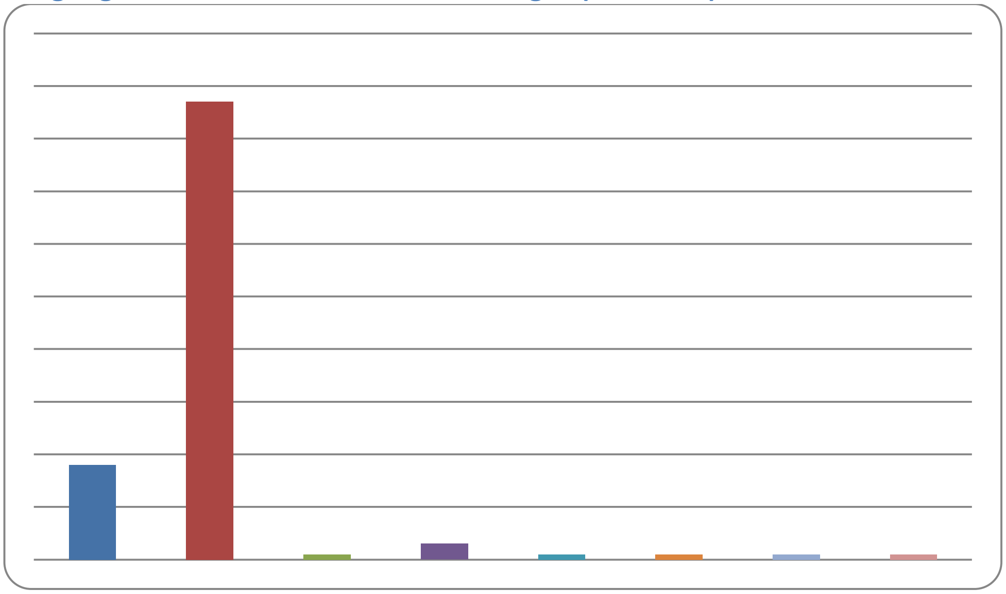
This question asks about the nature of the activity. The results show that most of them are educational programs.

Domain of study of the participants of the activity (Free text)



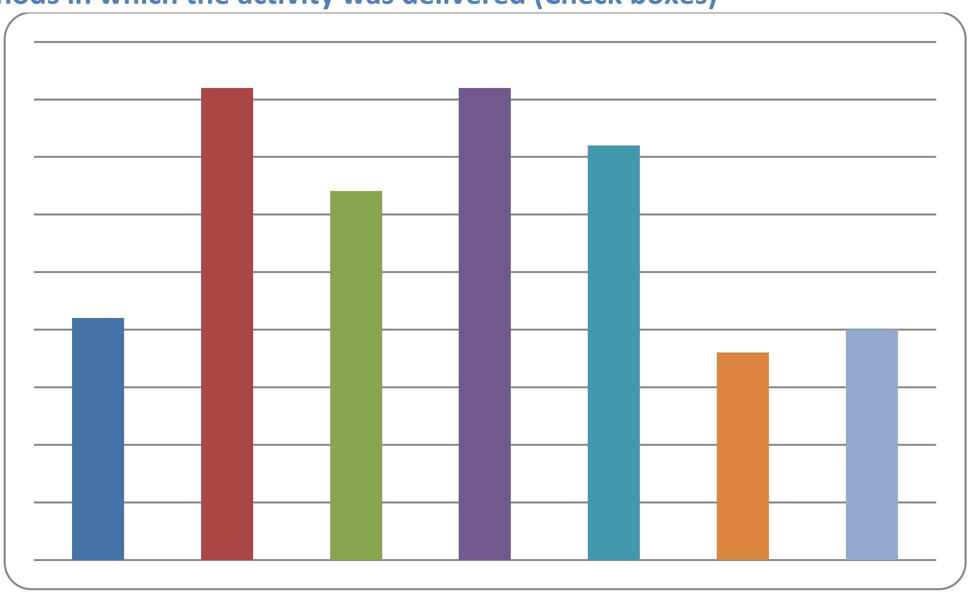
This question is trying to extract information on the domain of study of the participants of the survey. As expected, most activities are aimed at people in computer science or some other computational science, since HPC is directly related to them. However, it is also obvious that these activities also target physicists and mathematicians, both areas in which HPC could be highly beneficial.

Language in which the activities are taught (Free text)



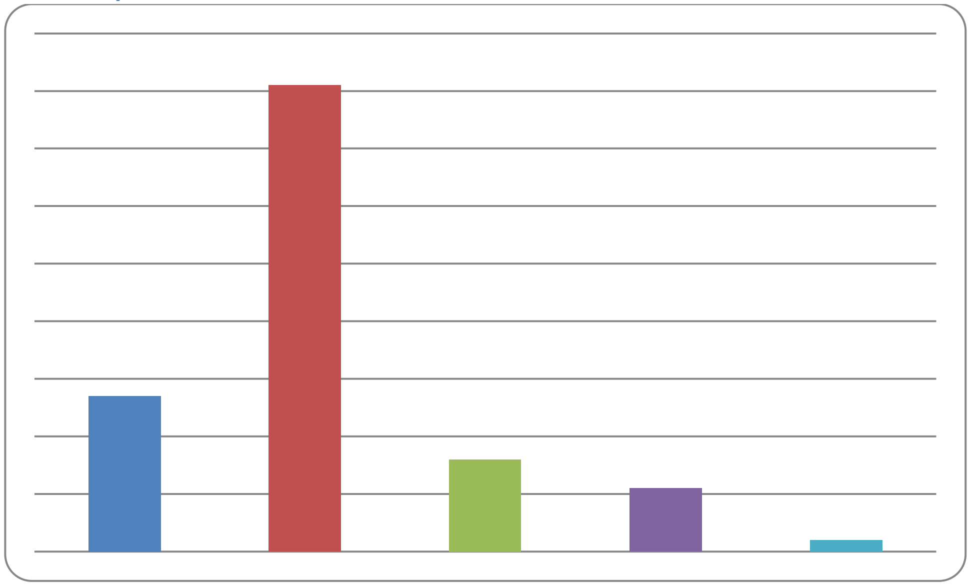
For the above bar chart, it is clear that most such activities are taught in English.

Methods in which the activity was delivered (Check boxes)



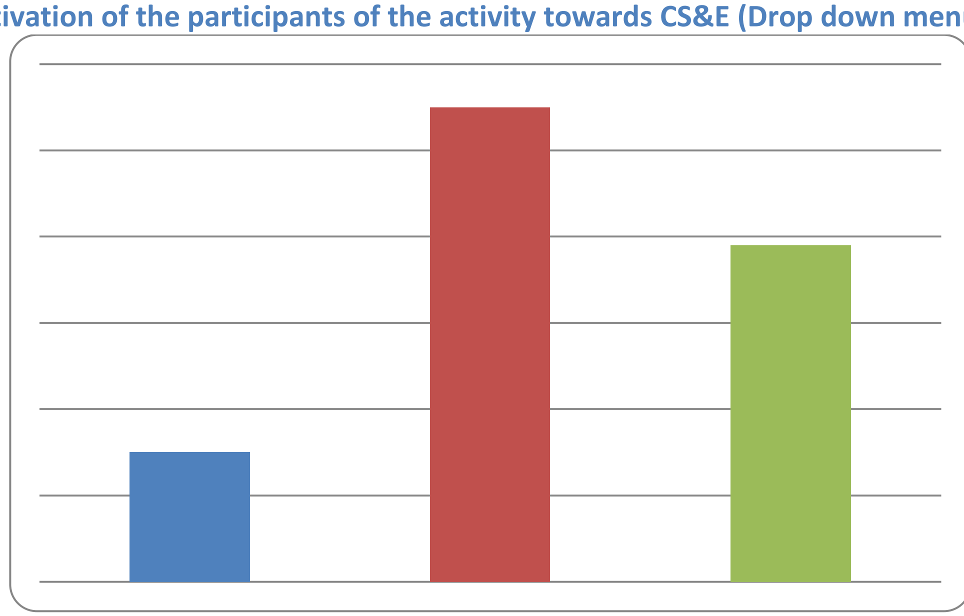
This question asks the respondents of the survey to state how the materials covered in their activities where delivered to their audience.

Methods used by the activity to interact with the participants of the activity (Check boxes)



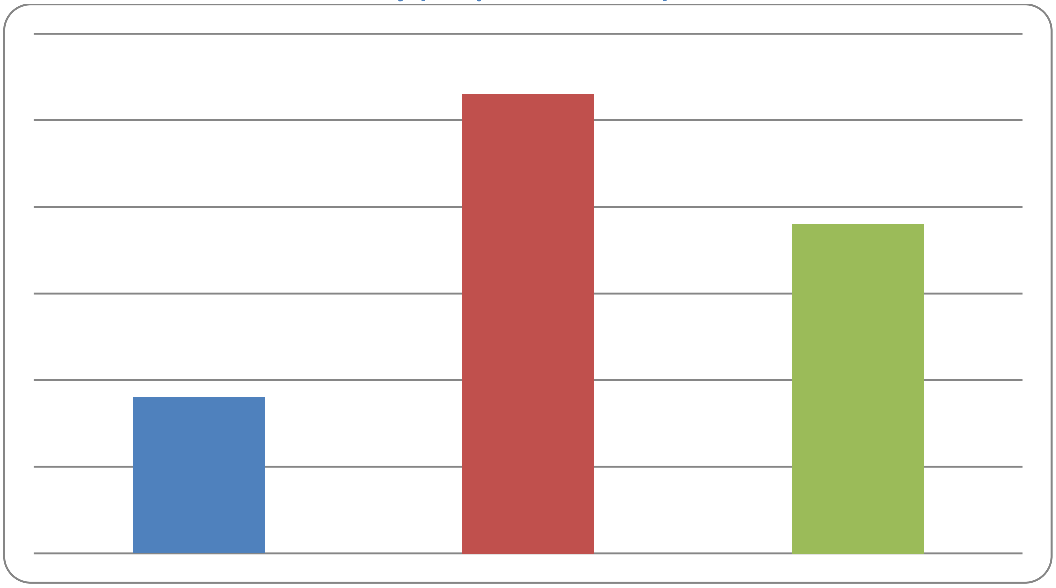
The above graph shows how the interaction between the activity and its participants took place. The results show that in most cases, the participation is live, which is usually the most common method of delivering such informational activities.

Motivation of the participants of the activity towards CS&E (Drop down menu)



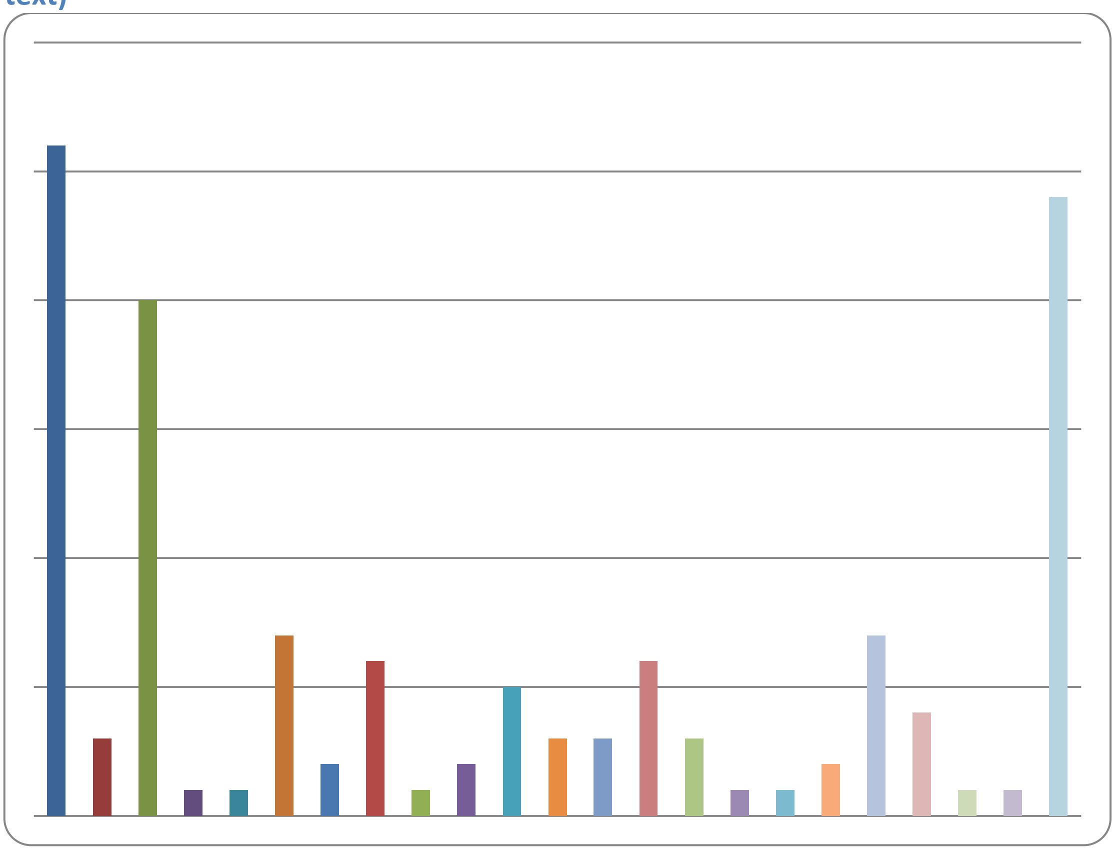
This question asks the respondent of the survey whether the audience targeted by the activity has already decided to follow HPC/CS&E, or whether they had not yet.

Selection criteria for the activity (Drop-down menu)



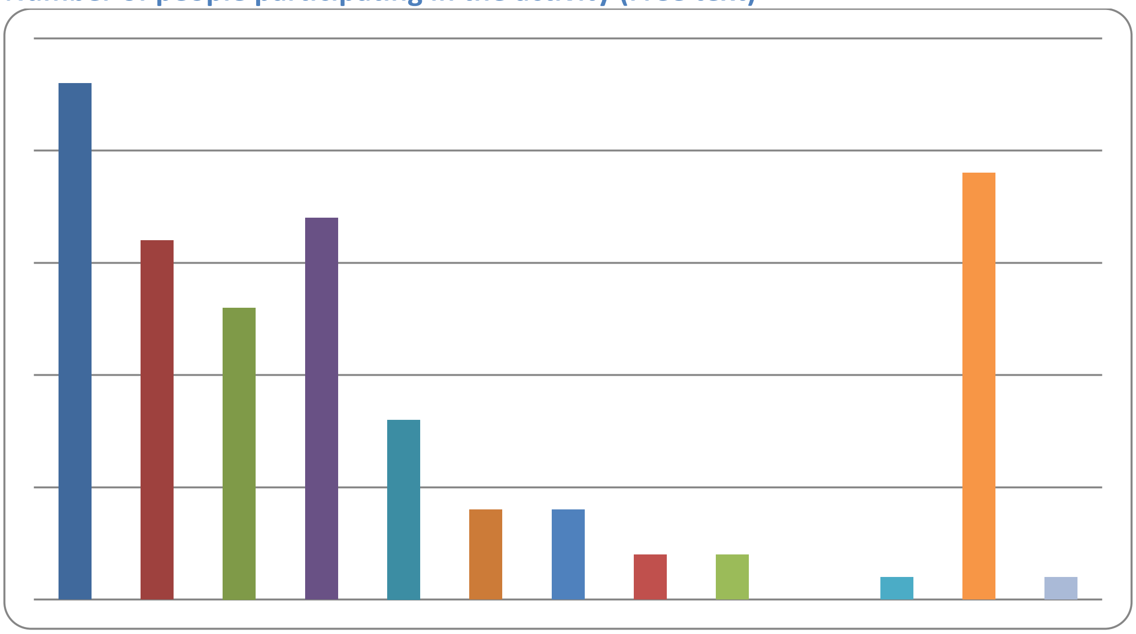
The above graph shows that most of the activities had some criteria in order to allow people to take part in it.

Minimum qualifications needed by the participant to follow the activity (Free text)



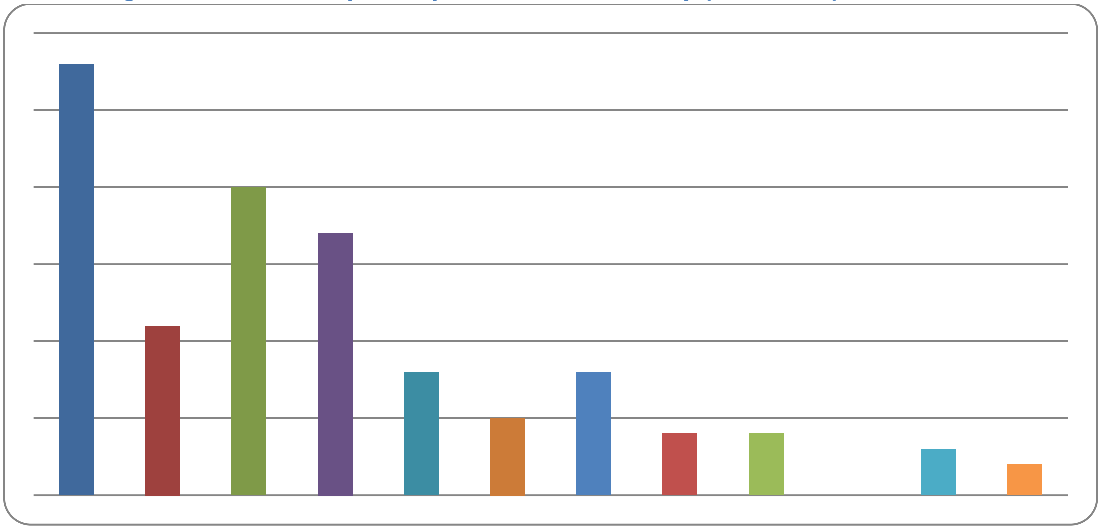
The question asks whether the activity has any prerequisite skills. From the results of our survey, it seems that in order to follow most activities, participants should have some knowledge in programming. This seems quite obvious since HPC is the high end of computational sciences which almost always involve programming.

Number of people participating in the activity (Free text)



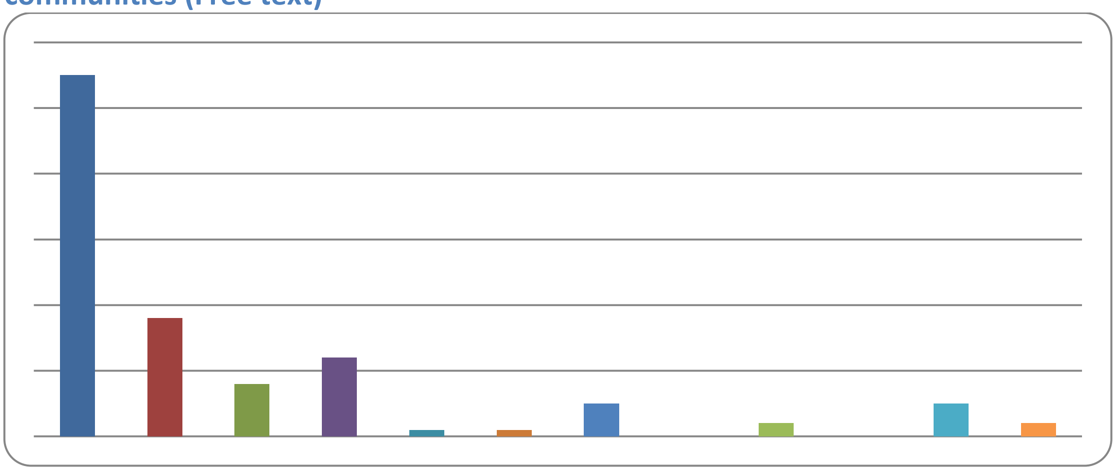
This question is asking for the number of people participating in these activities. The above results show a large number of activities with over one hundred participants. These activities may likely correspond to conferences or large classes (maybe curriculum classes which usually have many students attending). On the other hand, the graph also shows that smaller activities (with less than thirty participants), are also common. Such cases probably refer to workshops, labs, open days and smaller classes.

Percentage of the female participants of the activity (Free text)



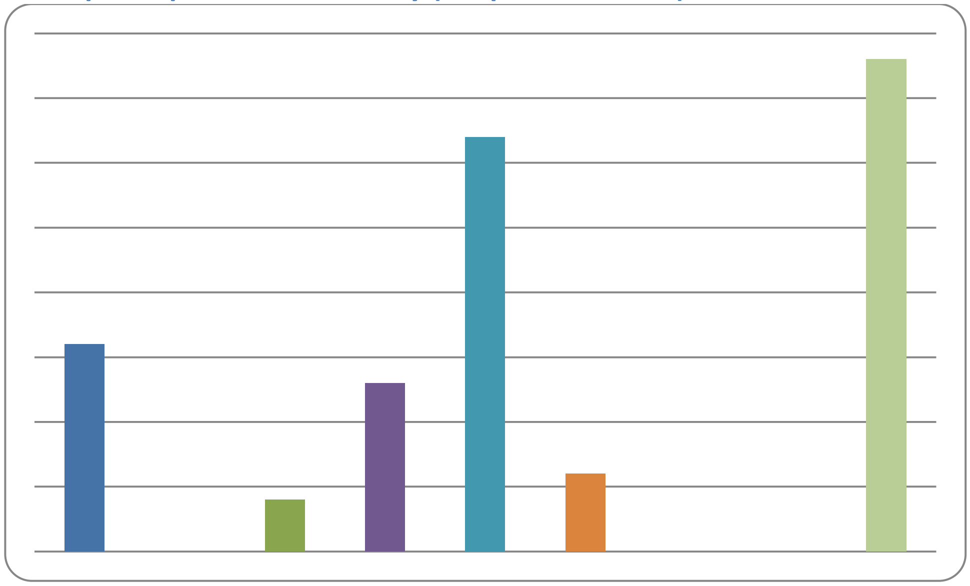
This question is asking for the percentage of females occupying the participants of the survey. The results show that usually, there is a significantly less number of women participants compared to men participants. This result was expected since many surveys prior to ours show that the percentage of females involved in computer science related topics is limited compared with the percentage of males.

Percentage of the participants of the activity from underrepresented communities (Free text)



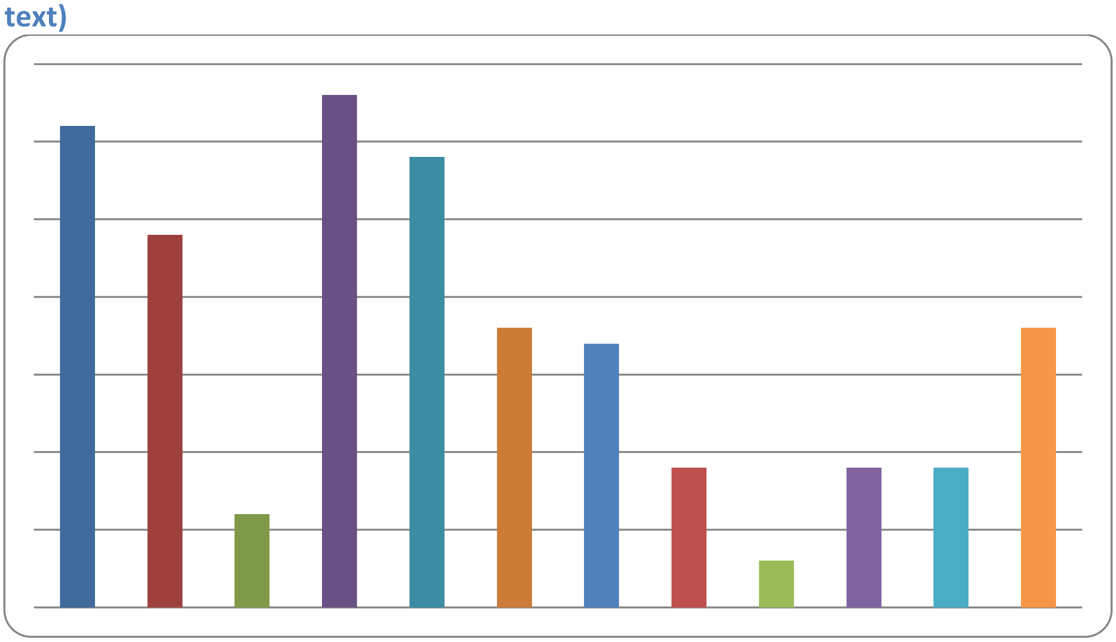
The graph shows how many people from underrepresented communities participate in the activity. Most people have either ignored the question or were unfamiliar with the answer (extremely high number of unspecified answers), and so it is difficult to reach to any conclusions.

Age of the participants of the survey (Drop-down menu)



The graph shows the age of the participants of the activity. If we ignore the "mixed" category which does not tell us much, we see that most of the participants are between the 19 years old and 24 years old. This is around the average age in which people are studying at a university or college.

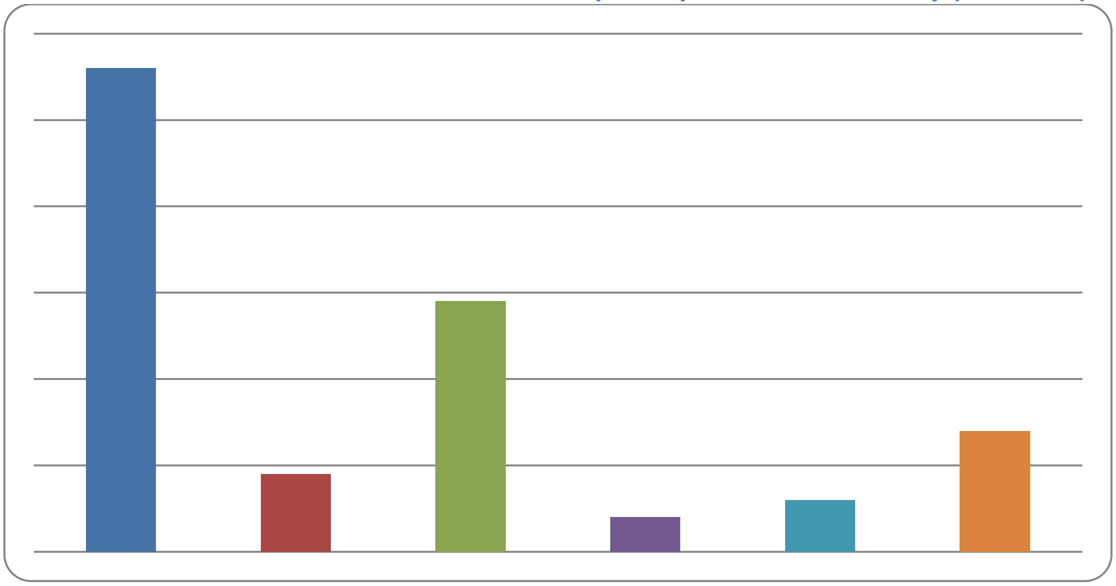
Preferred methods of the person filling the survey for promoting his activity (Free text)



The graph shows how the respondents of the survey choose to promote their activities. The results of the survey indicate that people usually use multiple promotion methods. Of course, sending an email seems to be the most popular, since it is very easy and it can be easily sent to many people simultaneously.

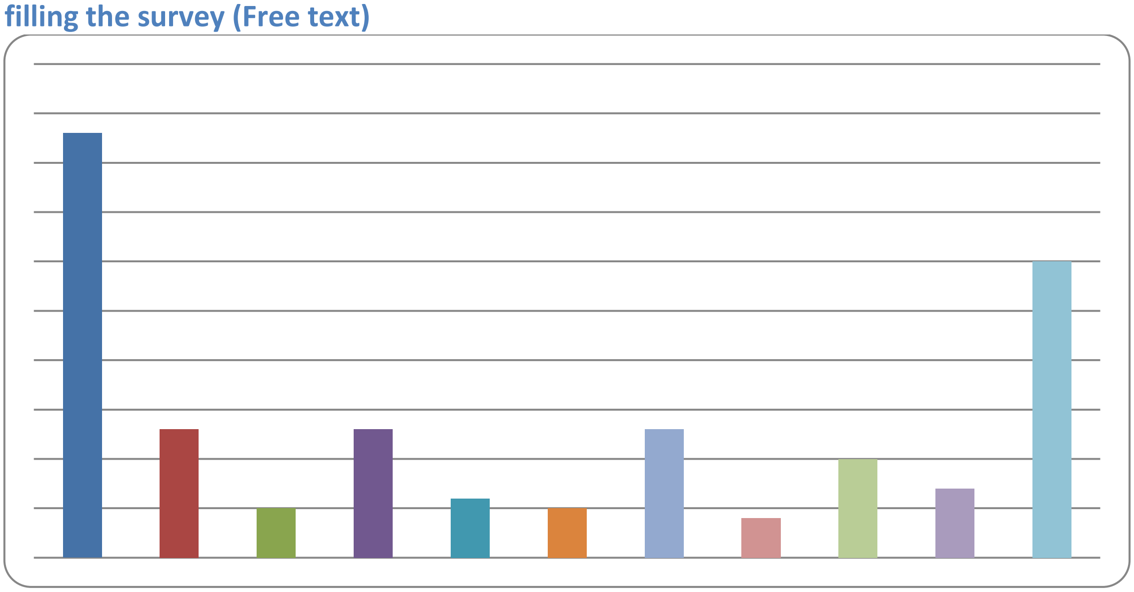
However, the WEB also seems to be a very common and interesting promotion method.

Quantitative measures used to evaluate the participants of the activity (Free text)



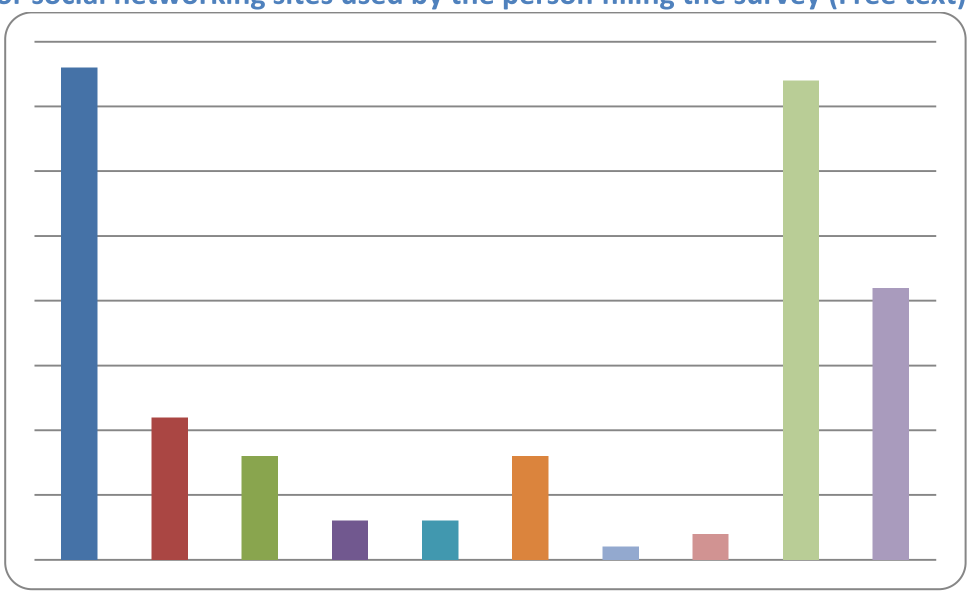
This question requires the respondents of the survey to state their evaluation methods, based on quantitative measures. The most common such measure seems to be student evaluation forms (e.g. survey). In some cases – most probably when the activity is actually part of the curriculum of an institute – the grade of the students was used as an evaluation measure. Some went even further and measured the success of their activity by monitoring the career of its participants.

Features that contributed to the success of the activity according to the person filling the survey (Free text)



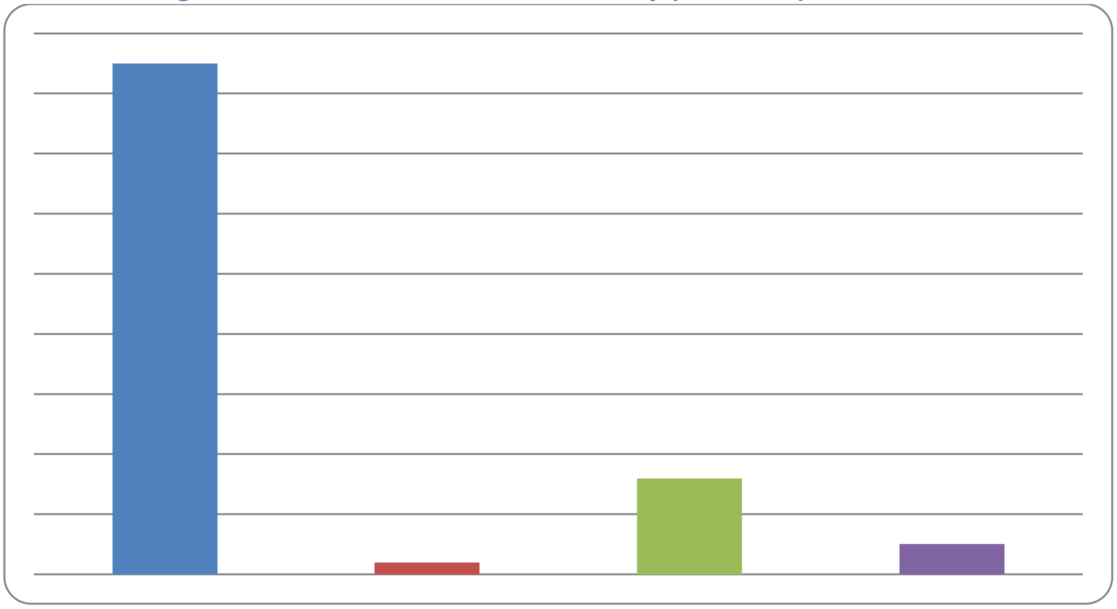
This question received many diverse answers (hence the large number of answers located in the “Other” category). However, we managed to locate some features which were common in many of the answers received. For example, many of the respondents of the survey believe that the practical involvement of the participants in the activity (“hands-on”) was a major feature that made the activity more interesting and successful. According the results, another such feature is the good teaching style used during the activity, as well as the quality of the staff helping with the activity and the students involved. Some also believe that the content of the activity was also a factor. Of course, for cases where the activity is part of the curriculum, the success of the activity could also have been due to the fact that students worked hard in order to get a good grade.

CSE or social networking sites used by the person filling the survey (Free text)



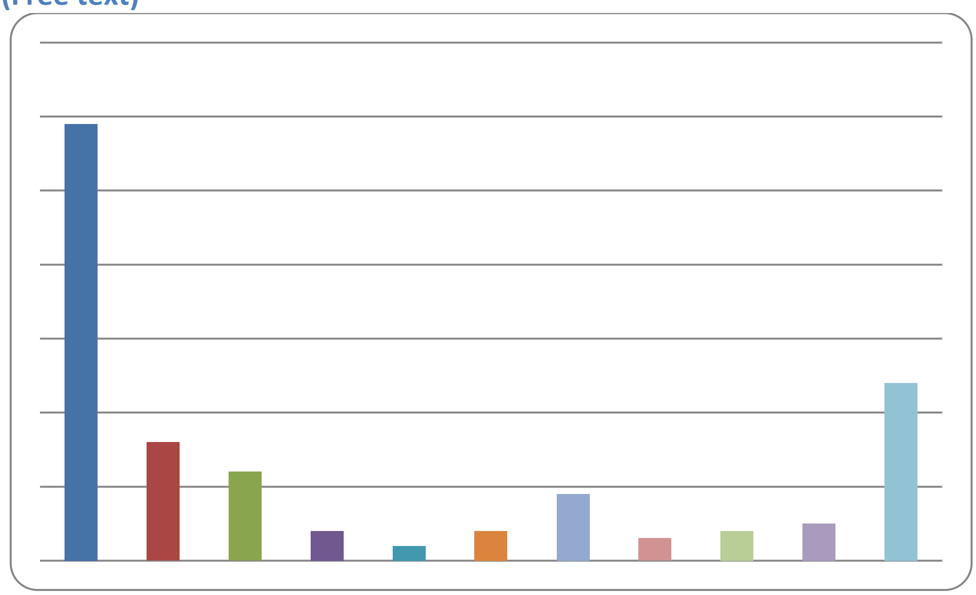
This bar chart shows the main CS&E or social networking sites in which the person filling in our survey is part of. The results seem to indicate that many of our respondents do not belong in any such communities. This of course could be because they do not think that being part of such networks would benefit them at all. On the other hand however, it might indicate a lack of such communities.

WEB functionality which, according to the person completing the survey can enhance the generation of a CSE WEB community (Free text)



This question is asking for the respondents' opinion as to what WEB functionality would help in improving the CS&E community. Most of them did not fill in an answer, which means that many have not probably even thought about this issue.

WEB sites used by the person filling the survey that help him stay informed about CSE (Free text)



This question is trying to collect the sites most commonly used by the respondents of our survey in order for them to stay informed about HPC and CS&E.

Appendix B: Overview of Partner Policy Reports

German Research School for Simulation Sciences

The German Research School for Simulation Sciences (GRS) provides the facilities and environment for interdisciplinary research and education in the applications and methods of HPC-driven simulation in science and engineering, next door to one of the largest German supercomputing centers, JSC.

GRS has four laboratories. The first focuses on simulations of large technical applications, in particular in the field of fluid dynamics. The second focuses on computational molecular biology. The third does research in the field of strongly correlated systems with the help of massively parallel computer simulations. Finally, the fourth develops tools that support simulation scientists in exploiting parallelism at massive scales.

Each laboratory is involved in the GRS Master's program "Simulation Sciences" aiming to enable the next generation of computational scientists and engineers to make efficient use of simulation on modern high-performance computers for their research, as well as a doctoral program aiming to give excellent graduate students from all over the world the opportunity to participate in strongly-interdisciplinary research in the fast evolving field of simulation sciences.

Both the master's and the doctoral program are publicized and disseminated via the following channels:

- The GRS WEB-site www.grs-sim.de,
- Entries in WEB portals such as www.studieren.de and www.daad.de,
- Poster and brochure mailings to other German and international universities,
- Advertisements in student magazines and newspapers,
- Student job fairs,
- Other special events (i.e. the official openings of the buildings in Aachen and Jülich.

Of these promotion methods, the WEB portal entries, especially the one on www.daad.de, appear to be the most successful.

Jülich Supercomputing Centre

The Jülich Supercomputing Centre offers three programs in an attempt to attract youth into computational sciences and high performance computing.

The first such program is the "JSC Guest Student Programme on Scientific Computing", which is held annually and goes on for ten weeks. The general goal of the program is to attract young scientists to HPC and CS, and to boost cooperation with universities. The program starts with an introductory course concerning the techniques of parallel computing and use of the Jülich supercomputers. Each student is then assigned a task according to his/her current field of work. At the end, each student has to prepare a report on his/her work. They also have to give a presentation on their work and discuss their results with other students.

The second activity held at the JSC, is called "Der Supercomputer als Teleskop – Die Sternenbewegung im Zeitraffer". The activity is basically an online contest in which the operation of a supercomputer is demonstrated and the interaction of stars and the resulting star dynamics are simulated to track the evolution of a galaxy. This activity addresses K-12 students (around 15-18 years old). About 130 such students participate in the activity for 6 weeks, learning about this interface between mathematics and supercomputing and experiencing the advantage of distributing and calculating work among a large number of groups. The fascinating topic of the program, as well as its competitive nature, was the source of great motivation from the students' perspective.

The final activity is called "SuperResi (Superrechner-Simulation)" and its intention is to give a first introduction into parallel processing to pupils at the age of about 14 years. The activity is designed around a program, which simulates a parallel computer by connecting several workstations via network to a cluster. Every workstation represents one CPU within the parallel computer. The participants create parallel solutions in pseudo code that solve given problems using one or more processors of the cluster. After a short introduction, the pupils participate in hands-on activities and experience quite easily what it means to work with parallel computers and how their thinking has to change in order to parallelize a problem. In some cases, the activity took the form of a competition to further motivate the students.

There are a number of ways in which JSC promotes itself and its activities to the outside world:

- Jülich computers are now occupying leading positions in the TOP500 (www.top500.org), the global ranking of the world's fastest supercomputers. This fact has a great impact on the public awareness and also on the attraction of students that are interested in the fields of computational sciences and engineering
- JSC gives its activities a competitive character (sometimes offering rewards to the winners). The success of these activities is shown by the constant requests for internships or visits by school classes
- JSC regularly participates in relevant science fairs as an attempt to promote their activities to young people.

Bergische Universität Wuppertal (BUW)

The BUW has a long tradition in high performance computing with local supercomputing facilities dedicated to research in particle physics and applied mathematics. In 2004, it also managed to acquire the supercomputer ALICEnext, which was ranked in the 74^{th} place in the TOP500. In the same year, a master's program was created, called "Computer Simulation in Science" (CSiS), which was a pioneering step in Germany.

In order to increase public awareness of CS and HPC related topics at BUW, it was very important to start good personal and scientific relations with other European institutes. Such was the case with Dr. Mike Peardon at Trinity College in Dublin, who is in charge of the Master's program in HPC there. Together, a joint seminar is organized on applications of HPC in natural sciences for students doing the master's in CSiS at BUW and students doing the master's in HPC at Trinity College Dublin, Ireland. The seminar is broadcasted by a video-conference link. During the seminar, users who are expert in HPC describe and discuss their research projects.

Another activity held to improve the outreach of CS and HPC at BUW is an exchange program of Japanese female students from Ochanomizu University in Tokyo, who come to BUW for one semester to attend classes in CSiS as well as in Chemistry and Physics. At the end of their stay, Japanese students do examinations of their courses at BUW and earn credit points recognized by the Ochanomizu University.

Finally, professors from the BUW visit local high schools in an attempt to attract the younger generation into CS and HPC and make them aware of the activities held in BUW.

National Center for Supercomputing Applications (NCSA)

The NCSA provides powerful computers and expert support that help thousands of scientists and engineers across the country to improve our world. It provides its available computational power to researchers, so that they can simulate all kinds of complex natural and engineered phenomena such as how galaxies collide and merge, how proteins fold, how molecules move through the wall of a cell, how tornadoes and hurricanes form and other phenomena.

One of the main tasks of NCSA is to prepare the next generation of scientists and engineers to make effective use of the computational tools and techniques available to them. In an attempt to achieve this task, it has launched three activities:

The first such activity is called "Virtual School of Computational Science and Engineering (VSCSE)". Its aim is to help graduate students, post-docs and young professionals from all disciplines and institutions across the country gain the skills they need to use advanced computational resources to advance their research. The courses are delivered simultaneously at multiple locations across the US using high-definition videoconferencing technology and via on-line WEB-based technologies to people at their own desktop.

The second activity carried out by the NCSA, is called "Institute for Chemistry Literacy through Computational Science (ICLCS)". This is a 5-year program to increase the chemistry literacy and chemistry-related pedagogical skills of rural

Illinois high school teachers. The goal of the program is to prepare these teachers and their students for the 21st Century through content, computational tools, teaching methodology and leadership development to meet the following goals:

- Strengthen high school teachers' and students' understanding of chemistry and the application of chemistry to the world around them
- Instill in teachers a sense of confidence and competence about their ability to teach chemistry, with a focus on using computational tools, modeling and visualization
- Build a strong learning community among research faculty and high school teachers to enable year-round professional development
- Create a cadre of leaders who will become advocates for excellence in mathematics and science.

This activity has been statistically proven to improve the teaching methods of Chemistry in high schools.

The last activity carried out by the NCSA is called "CyberInfrastructure Tutor (CI-Tutor)", which is a WEB-based training site for high performance computing and cyberinfrastructure topics. The CI-Tutor offers over 26 on-line tutorials, including the following that focus on CS&E:

- BigSim: Simulating PetaFLOPS Supercomputers
- Debugging Serial and Parallel Codes
- Getting Started on the TeraGrid
- Intermediate MPI
- Introduction to MPI
- Introduction to Multi-core Performance
- Introduction to OpenMP
- Introduction to Visualization
- Multilevel Parallel Programming
- Parallel Computing Explained
- Parallel Numerical Libraries
- Performance Tuning for Clusters
- Tuning Applications for High Performance Networks

The materials have been used by over 47,000 people over the last 10 years. The materials have been used in high school computer science courses, and have recently been incorporated into undergraduate classroom course materials at various universities.

The NCSA activities are publicized in the following ways:

- Word of mouth, which is a very effective means of recruiting people, but is also a process which is slow and limited in scope
- Through its WEBsite

- Through a regular newsletter
- Through connections with representative on campuses across the nation

The NCSA is also pioneering efforts to engage non-traditional communities of HPC/CST practitioners such as people from humanities, arts and social science communities.

The NCSA sees a need to attract more faculty and teachers engaged in pre-service education, which will indirectly increase the interest in CST through these well-trained teachers.

Another issue the NCSA would like to address is the small percentages of under-represented populations attending their activities. The reason for this is that their institutions/organizations do not usually provide funding for professional development efforts.

Shodor Education Foundation

The Shodor Education Foundation is an organization solely focusing on improving education in science, mathematics and medicine *through the effective use of simulations, modeling and more general computer science*.

Their goal is achieved via "Computing Matters": an array of activities targeting a wide audience:

- It excites elementary and middle school students and helps them build an interest in science, mathematics, engineering and technology,
- It illustrates career opportunities to high school students.

The curriculum, which has been aligned to U.S. national standards, includes the following four areas:

1. A general skills area, which includes the very basics such as math and verbal skills, office ethics etc.
2. A "Science Modeling" area, which includes an introduction to agent and system modeling, as well as an introduction to the basics of spreadsheets.
3. A "Graphics and WEB Design" area in which students are introduced to HTML, CSS, GIMP, Inkscape etc.
4. A "Programming and System Administrator" area which teaches the students the basics of programming, system administration (using the Unix command line, scripting, data management etc.) etc.