Supporting Statement (3145-0136)
Request For Clearance: National Science Foundation, Directorate of Education and Human Resources, Division of Graduate Education Distance Monitoring Data Collection in Support of the Computer & Information Science & Engineering's Pathways to a Revitalized Undergraduate Computing Education Program (CPATH)

Section A

Introduction

This request for Office of Management and Budget (OMB) review asks for clearance for a revised distance monitoring data collection in support of the Computer & Information Science & Engineering's Pathways to a Revitalized Undergraduate Computing Education Program (CPATH), which is administered by the National Science Foundation (NSF)'s Directorate for Computer and Information Science and Engineering (CISE). This clearance is requested under the Directorate for Education and Human Resources (EHR) Generic Clearance, OMB 3145-0136, which expires on March 31, 2011. The EHR Generic Clearance includes collections of information about NSF's education and training (E&T) activities. The survey was originally cleared under the generic clearance last year, but the initial data collection suggested several changes that have been incorporated in the revised survey. Namely, this version of the survey has incorporated numerous skip patterns to customize the survey experience. These changes were made to reduce the burden on PIs who because of project variety, may not need to respond to several sets of questions. As a result, the new monitoring survey will collect more accurate and relevant data.

A.1. Circumstances Requiring the Collection of Data

As part of its mission CISE contributes to the education and training of the next generation of computer scientists and engineers. Through the CPATH program, CISE is challenging its community partners – colleges, universities, and other stakeholders committed to advancing the field of computing and its impact – to transform undergraduate computing education on a national scale in order to meet the challenges and opportunities of a world where computing is essential to U.S. leadership and economic competitiveness across all sectors of society.

The use of computers has permeated and in many cases transformed almost all aspects of our everyday lives. As computer use becomes more important in all sectors of society, so does the need for preparation of a globally competitive U.S. workforce with knowledge and understanding of critical computing concepts, methodologies, and techniques. Thus, upgrading undergraduate computing education to keep abreast of the multitudes of rapid changes in computing is paramount for the U.S. economy and competitiveness.

The CPATH vision is of a U.S. workforce with the computing competencies and skills imperative to the Nation's health, security, and prosperity in the 21st century. This workforce includes a cadre of computing professionals prepared to contribute to sustained U.S. leadership in computing in a wide range of application domains and career fields, and a broader professional workforce with knowledge and understanding of critical computing concepts, methodologies, and techniques.

To achieve this vision, CISE is calling for colleges and universities to work together and with other stakeholders (industry, professional societies, and other types of organizations) to formulate and implement plans to revitalize undergraduate computing education in the United States. The full engagement of faculty and other individuals in CISE disciplines will be critical to success. Common challenges are fluctuating enrollments in traditional computer science programs, changes and trends in workforce demographics, the imperative to integrate fast-paced computing innovations into the curriculum, and the need to integrate computing concepts and methodologies into the undergraduate curriculum at large. Goals and strategies must be developed to address these and other challenges. Successful CPATH projects will be systemic in nature, address a broad range of issues, and have significant potential to contribute to the transformation and revitalization of undergraduate computing education on a national scale.

NSF needs data about CPATH awards for project and program monitoring to fulfill policy and program reporting needs. The data collected as part of OMB 3145-0136 will allow NSF officials to document the overall program investment in individual alliances and make future program policy decisions.

The purpose of the revised distance monitoring system is to provide updated data in each subsequent year of the project to compare to the baseline data collected with the original instrument as well as some new data that will serve as a baseline for future years. The CPATH PIs will be responsible for completing the data collection forms in the distance monitoring system.

The distance monitoring data collection form is provided in two forms: a text version in Appendix A and screen-shots from the on-line version in Appendix B. Appendix C contains a cross-walk between the questions on the monitoring survey and the EHR generic clearance tasks. Finally, Appendix D contains the e-mails that will be sent to PIs in various stages of the data collection.

A.2. Purposes and Uses of the Data

The information to be collected in this online system is required for effective administration, communication, and program and project monitoring; for meeting reporting requirements; for measuring the implementation of this program, its projects, and strategic goals as laid out in NSF's Strategic Plan; and for establishing a baseline for future program evaluations.

The primary purpose of this collection is program management, also known as program monitoring. This data collection activity is intended to track the design of the interventions, the strategies implemented, and the stakeholders engaged by CPATH awards to meet the objectives of the program. The information will be used by the CISE Directorate to administer and monitor the progress of the CPATH program. The findings will be used to recommend, among other things, administrative changes in program specifications, individual program focus and emphasis, and recruiting efforts.

In recent guidance from the Director of OMB, M-10-32, the need for rigorous evaluations and the objectives of evaluations of programs were clearly outlined including the use of evaluation resources. Because the collection of data contained in this monitoring effort contributes to the formal evaluation of the program and provides regular measures of program performance by accumulating operating information from each project in the program this guidance is particularly pertinent to this request.

The CPATH program will also use the data to fulfill reporting requirements. As part of its performance assessment activities, NSF relies on the judgment of external experts to maintain high standards of program management. Directorate and Office advisory committees (ACs) meet twice a year, while Committees of Visitors (COVs) for divisions or programs meet once every three years. Data collected in the CPATH monitoring system may be used to report to these committees on program activities.

In addition, NSF is required to measure the attainment of its program, project and strategic goals, as identified by the President's Accountable Government Initiative, the Government Performance and Results Act (GPRA) of 1993, and NSF's Strategic Plan. These data help NSF management examine their progress towards these goals and respond to these reporting requirements. The data will also be used to inform the program evaluation that SRI International is conducting.

Data are collected using a distance monitoring system, an online system developed to facilitate electronic data collection. Respondents will be all PIs of active CPATH awards. The survey will contain questions related to staff and project participant characteristics, project implementation characteristics, and project outputs. Appendix C contains a detailed list of question topics and how they relate to the EHR generic clearance tasks.

A.3. Use of Information Technology to Reduce Burden

The CPATH data collection will employ a Web-based system to facilitate respondents' data entry across computer platforms. One innovative feature of Web systems is the thorough checking of all submitted data for completeness, validity, and consistency. Checks are performed as data are entered. Most invalid data cannot enter the system, and questionable or incomplete entries are called to respondents' attention before they are submitted to NSF. The CPATH distance monitoring system has user-friendly features such as automated tabulation, data entry with custom controls such as checkboxes, data verification with error messages for easy online correction, standard menus, and predefined charts and graphics. All of these features facilitate the reporting process, provide useful and rapid feedback to the data providers, and reduce burden.

The data for this monitoring effort will be collected using 508-compliant Web-based forms. The question format will be primarily quick-response checkboxes, with number and text boxes provided for the addition of specific data or outstanding examples. Respondents may save, leave and re-enter their data collection form as often as they desire and continue to change their responses until they submit their form. Additionally, since the collection is Web-based, minor changes in wording and displays can be easily made in response to user feedback. This year, several skip patterns have been added to customize the survey and ensure PIs only respond to questions that are applicable to the focus of their project. The crosswalk (appendix C) describes the skip patterns included in the monitoring survey.

A.4. Efforts to Identify Duplication

This data collection will not duplicate other NSF efforts; comparable data are not collected on an annual basis for the CPATH program. The CPATH data collection will be coordinated with the NSF FastLane Project Reports system (OMB 3145-0058) to ensure that the two collections do not gather similar data. As much as possible, data from CPATH proposals and annual reports will be used to reduce the overall response burden. Additionally, aggregate data will be shared

with NSF-funded researchers as appropriate, thereby minimizing the possibility that other researchers will duplicate these efforts in their own future collections.

A.5. Small Business

No information is to be collected from small businesses.

A.6. Consequences of Not Collecting the Information

First, without this information NSF will be unable to document the effectiveness or outcomes of the CPATH program. Second, without this information NSF will be unable to comply fully with the GPRA reporting requirements or with the congressional mandate that the Foundation evaluate its science, technology, engineering, and mathematics (STEM) education programs. Third, without this information NSF will be unable to disseminate information to other projects and institutions about successful approaches to transform undergraduate computing education on a national scale with the goal of meeting the challenges and opportunities of a knowledge society where computing is essential to economic competitiveness and social welfare.

A.7. Special Circumstances Justifying Inconsistencies with Guidelines in 5 CFR 1320.6

The data collection will comply with 5 CFR 1320.6.

A.8. Consultation Outside the Agency

The notice inviting comments on the EHR Generic Clearance (OMB 3145-0136) was published in the Federal Register August 24, 2007, Volume 72, Number 164, page 48694. No comments were received.

The comments by CPATH PIs after the first data collection informed the changes that have been made to the monitoring survey. Former PIs were asked to pilot the monitoring survey and modifications were made based on their feedback.

A.9. Payments or Gifts to Respondents

No payments or gifts will be provided to respondents.

A.10. Assurance of Confidentiality

The data collected will be available only to the respondents, NSF, and members of SRI International's CPATH evaluation project team. Data will be processed according to Federal and State privacy statutes. To protect privacy, only composite data or graphical representations will be released to the public.

When respondents are presented with the first screen of the data collection form, they will be instructed as follows: "Information from this data collection system will be retained by the National Science Foundation, a federal agency, and will be an integral part of its Privacy Act System of Records in accordance with the Privacy Act of 1974 and maintained in the Education and Training System of Records 63 Fed. Reg. 264, 272 (January 5, 1998). These are confidential files accessible only to appropriate National Science Foundation (NSF) officials, their staffs, and their contractors responsible for monitoring, assessing, and evaluating NSF programs. Only data in highly aggregated form, or data explicitly requested as "for general use," will be made available to anyone outside of the National Science Foundation for research purposes. Data submitted will be used in accordance with criteria established by NSF for monitoring research and education grants, and in response to Public Law 99-383 and 42 USC 1885c."

A.11. Questions of a Sensitive Nature

There are no questions of a sensitive nature included in this data collection.

A.12 Estimates of Response Burden

A.12.1. Number of Respondents, Frequency of Response, and Annual Hour Burden

The Web-based collection will be an annual activity of the CPATH program. Currently there are 76 active CPATH awards and sites. The PIs on these awards will be required to submit data to the CPATH data collection system every year. Because new awards are likely to be added and some earlier awards will end, it is expected that the annual number of respondents will change somewhat over time.

The total number of annual person-hours for the first year of data collection is estimated at 228 hours. This burden is computed by multiplying the number of anticipated respondents (76) by the estimated response burden per person, based on pretesting of the data collection form by a sample of PIs.

Respondent Type	Number of Respondents	Burden Hours Per Respondent	Annual Person-Hours
PIs	76	3	228

A.12.2. Hour Burden Estimates by Each Form and Aggregate Hour Burdens

There is only one data collection form. The estimated annual burden is presented in section A.12.1 above.

A.12.3. Estimates of Annualized Cost to Respondents for the Hour Burdens

The overall annualized cost to the respondents is estimated to be $8,773.44.

Respondent Type	Number of Respondents	Burden Hours Per Respondent	Average Hourly Rate	Total Annual Costs (rounded)
PIs	76	3	$38.48	$8,773.44

The estimated hourly rate for PIs is based on an annual survey conducted by the American Association of University Professors, data from which were published in *The 2008-09 Chronicle Almanac*, 55(1): 24, accessed at http://chronicle.com. According to results of this survey, the average salary of an associate professor across all types of doctorate-granting institutions (public, private, church-related) is $80,043 for the 2008–09 academic year. This calculates to an average hourly rate of $38.48 when divided by the 2,080 hours in a standard work year.

A.13. Estimate of Total Capital and Startup Costs/Operation and Maintenance Costs to Respondents or Record Keepers

There will be no overall annual cost burden to respondents or record-keepers that results from the distance monitoring of the CPATH program other than the time spent responding to the data collection form.

It is usual and customary for individuals involved in education and training activities in the United States to keep descriptive records. The information being requested is from records that are maintained as part of normal educational or training practice. Furthermore, the majority of

PIs are active or former grantees or participants in programs or projects once funded by NSF. In order to be funded by NSF, institutions must follow the instructions in the NSF Grant Proposal Guide (GPG) that is cleared under OMB 3145-0058. The GPG requires that all applicants submit requests for NSF funding and that all active NSF awardees do administrative reporting via FastLane, an Internet-based forms system. Thus, the PIs who will be the primary respondents for the CPATH data collection will make use of standard office equipment (e.g., computers) and Internet connectivity that are already required as startup and maintenance costs under the NSF GPG.

A.14. Estimates of Costs to the Federal Government

The table below shows the average annual cost to NSF for the CPATH distance data collection, by operational activity. The developmental costs are listed as 0 as they were incurred in the initial design of the survey administration system. The remaining costs for this year as well as remaining option years, we estimate costs based on experience from administration of the monitoring survey in 2009. The total cost is allocated between the two remaining task groups based on past experience.

Operational Activity	Average Annual Cost
a. System development (includes initial development of the database and Web-based application and later changes requested by the program, e.g., increased reporting tools, additional validations)	-
b. System maintenance, updates, and technical support (updates required each year before opening the collection; maintenance required to keep the system current with technology, e.g., database servers, operating systems)	$67,500
c. Data collection opening and support (e.g., online and telephone support to respondents, contacting respondents to encourage completion of the questions); reporting as defined by the Division of Graduate Education; and follow-up activities (e.g., providing data to other consultants)	$22,500
Total Average Annual Cost to NSF	$90,000

A.15. Changes in Burden

The survey is longer than the version that was cleared in 2009; however, the survey has been modified to include skip patterns to better fit the wide variety of CPATH project types. Following feedback from some PIs that several questions in the last version of the monitoring survey did not apply to their project, questions have been added to categorize projects and ask more in depth questions only after the PI has confirmed that such questions are relevant to their project's area of focus. As such, pilot testing suggests the burden has not changed for the average project although for a few of the most complicated projects, the survey might take slightly longer than last year's version.

A.16. Plans for Publication, Analysis, and Schedule

Data collection is scheduled to begin in February 2011. The PIs will have 90 days to enter the data for their award site. Extensions will be granted by NSF program officers as necessary. Once the data collection has been completed, agency staff can access the data through the online system as needed.

Like many agencies, NSF is reducing its reliance on formal (i.e., traditional) publication methods and publication formats. SRI International, the contractor that manages the data collection Web site and database, is forbidden contractually from publishing results unless NSF has made a specific exception. In short, all products of the collections will be the property of NSF, and NSF will be the exclusive publisher of the information being gathered. Often it is only after seeing the quality of the information collected that NSF decides the format (raw or analytical) and manner (in the NSF-numbered product Online Document System (ODS) or simply a page on the NSF Web site) in which to publish.

The data from this collection will be used for internal review purposes, for monitoring the CPATH awards, and for reporting to Congress (e.g., the GPRA Annual Performance Plan). Reports to NSF management, PIs, and Congress dealing with characteristics and performance of the CPATH program will include statistical tables and charts generated from the database. The data from the monitoring survey was included in the first annual report and will also inform all future reports.

A.17. Approval to Not Display Expiration Date

Not applicable.

A.18. Exceptions to Item 19 of OMB Form 83-I

No exceptions apply.