CORE QUESTIONS and REPORT TEMPLATE
for
FY 2017 NSF COMMITTEE OF VISITOR (COV) REVIEWS

Guidance to NSF Staff: This document includes the FY 2017 set of Core Questions and the COV Report Template for use by NSF staff when preparing and conducting COVs during FY 2017. Specific guidance for NSF staff describing the COV review process is described in the "COV Reviews" section of NSF's Administrative Policies and Procedures which can be obtained at https://inside.nsf.gov/tools/toolsdocuments/Inside%20NSF%20Documents/Policy,%20Procedures,%20Roles%20and%20Responsibilities%20for%20COV%20Reviews%20and%20Program%20Portfolio%20Reviews.pdf[1].

NSF relies on the judgment of external experts to maintain high standards of program management, to provide advice for continuous improvement of NSF performance, and to ensure openness to the research and education community served by the Foundation. COV reviews provide NSF with external expert judgments in two areas: (1) assessments of the quality and integrity of program operations; and (2) program-level technical and managerial matters pertaining to proposal decisions.

The program(s) under review may include several sub-activities as well as NSF-wide activities. The directorate or division may instruct the COV to provide answers addressing a cluster or group of programs – a portfolio of activities integrated as a whole – or to provide answers specific to the sub-activities of the program, with the latter requiring more time but providing more detailed information.

The Division or Directorate may add questions relevant to the activities under review. Copies of the report template and the charge to the COV should be provided to OIA prior to forwarding to the COV. In order to provide COV members adequate time to read and consider the COV materials, including proposal jackets, COV members should be given access to the materials in the eJacket COV module approximately four weeks before the scheduled face-to-face meeting of the COV members. Before providing access to jackets, the Conflict of Interest and Confidentiality briefing for COV members should be conducted by webinar, during which, NSF staff should also summarize the scope of the program(s) under review and answer COV questions about the template.

Suggested sources of information for COVs to consider are provided for each item. As indicated, a resource for NSF staff preparing data for COVs is the Enterprise Information System (EIS) –Web COV module, which can be accessed by NSF staff only at http://budg-eis-01/eisportal/default.aspx. In addition, NSF staff preparing for the COV should consider other sources of information, as appropriate for the programs under review.

For programs using section IV (addressing portfolio balance), the program should provide the COV with a statement of the program's portfolio goals and ask specific questions about the program under review. Some suggestions regarding portfolio dimensions are given on the template. These suggestions will not be appropriate for all programs.

Guidance to the COV: The COV report should provide a balanced assessment of NSF's performance in the integrity and efficiency of the ***processes*** related to proposal review. Discussions leading to answers of the Core Questions will require study of confidential material such as declined proposals and reviewer comments. ***COV reports should not contain confidential material or specific information about declined proposals.*** The reports generated by COVs are made available to the public.

We encourage COV members to provide comments to NSF on how to improve in all areas, as well as suggestions for the COV process, format, and questions. For past COV reports, please see http://www.nsf.gov/od/oia/activities/cov/.

[1] This document has three parts: (1) Policy, (2) Procedures, and (3) Roles & Responsibilities.

Executive Summary

The Committee of Visitors (COV) for the Office of Advanced Cyberinfrastructure (OAC) met on November 28-30, 2017 at NSF as part of its review. This report summarizes the findings for that review along with several recommendations that the COV developed to address the findings.

The COV first heard presentations from the Office Director and Program Directors on the various programs that included information relevant to the review. All of the Program Directors were available to answer questions regarding the individual programs. The COV also met separately with senior leadership in OAC including the Associate Director for CISE. In all cases, the staff were forthcoming with data requested by the COV and responded to all the questions the committee had. Access to the eJacket system was also provided to the COV in order to review the individual grant proposals and analyses in order to perform our own assessment. In addition, OAC had provided materials prior to the actual meeting at NSF that included background information including the annual reports, prior COV reports, and other data deemed necessary for the review.

The result from this process includes a completed template provided below containing the details gathered through the review process and from discussions the COV held. This section, summarizes the major findings organized by the main components of the COV charge.

Assess the integrity, efficacy, and quality of the processes used to review, recommend, and document proposal actions.

The review analyses we had access to were thorough, thoughtful, and complete. In general, a good mix of reviewers with appropriate technical expertise were chosen. The COV finds that it would be useful to have a better sense of the breadth of reviewers (e.g. with respect to institution) in order to better gauge the integrity and quality of the process. The number of reviewers for each proposal was appropriate in most cases. However, the number of reviewers does not always align with the award size, where one would expect a larger number of reviewers for larger grant proposals. This sometimes appears to be dependent on individual programs as well.

In cases where the decision varies from the panel summaries, the Program Director (PD) provided good analyses for the final decision. Some jackets contained a wide range of ratings and in those cases the PD provided a good explanation/analysis for the variance that provided a fuller context for the final decision. The COV commends the PDs for exercising due diligence and extra effort in several cases to obtain clarifying information from PIs when necessary in making a decision.

The COV found that the terms used to rate proposals can be confusing (Responsive/Non-responsive/Highly responsive; Highly competitive, etc.). Some of the review analyses discussed budgets and some did not; again this sometimes appeared to be dependent of the program area (e.g. Data).

In some program areas (e.g. HPC) a large preponderance of reviewers were male, although it did not seem to have an adverse effect on overall proposal outcomes. The COV noted (through an informal analysis) that gender balance in the review panels varied widely across the program areas. We could not determine whether this was also true for URMs.

Finding. *The overall integrity, efficacy, and quality of the processes was excellent. In general, there was thorough documentation of all proposals with clear justification for the award decisions made.*

Recommendation. *Continue to maintain the high standards within the office and work to provide more consistency among the various program elements.*

All program review summaries included a category of broader impact, and most reviewers included it in their own assessments. However, it was not clear to the COV whether reviewers were selected for their ability to assess broader impact as a primary capability. For example, reviewers who specialized in education/training impact did not appear to be included in the broader merit discussion (outside the LWD program) even for virtual reviews. We observed several instances where a PD went back to a PI to ask for clarification or expansion of the broader impact statement if the panel reviews considered this a weakness (and where the intellectual merit/overall proposal rating was otherwise high). No instances were discovered in the jackets we read where a scientific proposal assessed with a very strong broader impact resulted in a dialog to increase the intellectual merit however. As well, it was not uniform practice for the PD to have a recorded dialog with a PI regarding the need to improve the broader impact.

Broader impact discussions sometimes appeared to include areas that the COV thought more accurately described intellectual merit, or in other cases was limited to mentoring of project participants. It is commendable, however, that scientific portals were considered broader impact.

Specific to the use of the broader impact criteria within the review process, the COV had the following findings and recommendations:

Finding. *In many jackets, the PD included a robust discussion of broader impact. The COV did note however that the quality of the broader impact discussion varied considerably – far more than the quality of the intellectual merit discussion. The COV also noted that in the discussions, the broader impact criteria appeared to be significantly secondary to the intellectual merit criteria – that is, broader impact might be used to distinguish among proposals having similar intellectual merit.*

Recommendation. *The COV believes it would be worthwhile to develop strategies to leverage track record in successful achievement of the broader impacts as part of the assessment in the review process, whether that track record is recorded qualitatively or quantitatively.*

Assess the quality of project management, monitoring, and evaluation of funded proposals

The program's responsiveness to emerging research and education opportunities is excellent. The Office uses a variety of strategies to learn about the needs both within the Foundation and the community at large. Examples of these types of activities include participation in NSF cross-directorate panels and working groups (HDR, Quantum Learning, Arctic) as well as external groups, including NITRD, and participation in interagency reviews. In addition, the PDs attend annual PI meetings and professional conferences in their fields to stay abreast of the latest trends within their fields.

OAC has a well-defined set of activities to help it plan and prioritize its various programs. In addition to the ones listed above, the Office also has regular meetings with its advisory committee (ACCI), which provides external advice on the programs as well as recommendations from various working groups on topics related to cyberinfrastructure. OAC thinks strategically when making decisions about proposals that will benefit the entire Foundation and not just their own Directorate as evidenced by several examples provided by senior leadership. The position of Science Advisor also plays another important role for OAC to learn about new strategic directions and opportunities for Cyberinfrastructure.

Annual reports and site visit assessments can play a critical role in the monitoring of awards. OAC should continue to use these judiciously.

Finding. *The overall management of the program appears to be excellent. The entire office is highly responsive to the rapidly changing environment and appears to make diligent efforts to both learn what the community needs are, and to plan and prioritize the activities within OAC to address those needs.*

Recommendation. *COV recommends that OAC (in its leadership role within cyberinfrastructure) investigate the use of more automated tools for monitoring program management.*

Finding. *The COV noted that dwell time appears to be improving, although it is still a concern.*

Recommendation. *We encourage OAC to do a fuller analysis to identify process bottlenecks and potential areas to modernize processes through automation and data sciences.*

Comment on the Office's balance, priorities, and strategies for realizing the potential of the Office, and any other issues you think are relevant to quality and integrity of the merit review process, including technical and managerial matters pertaining to proposal recommendations.

Finding. *COV was impressed with the great level of attention and detail that the Office has given to the large number of proposals that they must review. We commend OAC for seeking to ensure consistency in the review process despite having a large number of variables to deal with in the review process. Some suggestions for achieving greater consistency include:*

Recommendation. *OAC should consider including a sustainability plan for software and tools as a criterion in the review process in all programs where it is applicable. In addition, OAC might consider requiring certain awards to have a report with a discussion of community acceptance of the software for renewals.*

Recommendation. *OAC should have a well-defined process and a clear understanding of the research and industry landscape in the development of DCL and solicitations. This clear understanding is critical to the mission of OAC.*

Recommendation: *OAC should consider automation/deep learning + science as a theme meriting prioritization across the portfolio. In addition, it would be helpful if future annual reports include a discussion of the potentially transformational implications of deep learning on the conduct of science.*

Requests advice on progress concerning issues raised by the previous (2011) COV during the last four years, recognizing the change in administrative structure that took place in 2013 and the reassessment in 2016

The previous COV report had several recommendations with regard to the dwell time for some of the larger grants, the size and duration of the DataNet awards, the tracking of outcomes of the Taskforce reports from 2009, the skill sets required for the management of the large-scale programs such as Track 1 and DataNet, and the turnover in some of the program staff.

Finding. *In each case, OAC provided an appropriate and thoughtful response to the prior COV's findings and recommendations.*

Recommendation. *As many of the findings reflect the rapidly changing environment (both technically and fiscally) and, in some cases, longer-term projects, the COV encourages the Office to continue to closely monitor and address the issues highlighted in earlier reports. The COV also noted that dwell time continues to be a concern and OAC should continue to closely monitor this issue.*

As in the past, the technologies that underlie advanced cyberinfrastructure continue to evolve rapidly, presenting a persistent challenge to deliver stable platforms in the face of this volatility. Propelled by the greater abundance and affordability of computing, storage and network performance, a wider range of applications have embraced computational methods and scalable data analytics. Among the emerging communities shifting use away from the established incumbent users, the application of deep learning shows strong momentum. Responding to this trend, commercial vendors of both hardware and those hosting elastic environments (cloud) have significantly advanced the ability of their platforms to deliver high performance training for machine learning. The potential for machine perception and data analysis to broadly transform the practice of science itself is evident, placing OAC at the heart of this revolution. Thus, the OAC mission faces challenges in:

1. Maintaining expertise in advanced computing hardware
2. Sustaining software compatibility and scalability to fully leverage state of the art hardware
3. Advising both incumbent and newly-interested user communities in readiness of maturing technologies for desired applications as well as potential to overcome prior limitations

OAC is not alone in these challenges and could significantly benefit from alignment with other National Strategic Computing Initiative participants. Further, the commercial elastic infrastructure (cloud) in 2018 and beyond may offer hardware and software architecture significantly closer to NSF mission needs than in prior reviews. For example, we are aware of at least one major cloud provider's recent partnership with a supercomputing vendor to co-locate leadership-class architecture with elastic compute and storage.

The COV understands that OAC may already be in discussions on these and related topics and that the recently released National Academies Report, *Future Directions for NSF Advanced Computing Infrastructure to Support U.S. Science and Engineering in 2017-2020,* has also provided several recommendations. We urge OAC to continue to investigate these trends and to incorporate them into their annual reports.

FY 2017 REPORT TEMPLATE FOR NSF COMMITTEES OF VISITORS (COVs)

The table below should be completed by program staff. *Note: In the following standard report template, NSF provides guidance concerning where relevant information can be found in the foundation's administrative data systems. To facilitate the work of the COV, CISE prepared annual reports based on these data sets, eliminating the need for the panel members to pull the data themselves. Annual reports for the period under study have been provided on the COV website.*

Date of COV:	November 28-30, 2017
Program/Cluster/Section:	CISE/OAC
Division:	Office of Advanced Cyberinfrastructure (OAC)
Directorate:	Computer and Information Science and Engineering (CISE)

Number of actions reviewed:

Year	Awards	Declinations	Other
2013	40	67	N/A
2014	36	27	N/A
2015	50	50	N/A
2016	41	36	N/A

Total number of actions within Program/Cluster/Division during period under review:

Year	Awards	Declinations	Other
2013	139	349	N/A
2014	138	202	N/A
2015	176	295	N/A
2016	179	309	N/A

Other: *Proposals in programs managed by divisions or directorates outside of OAC (such as Big Data Science & Engineering proposals) were excluded from the random sampling, as were proposals that belong to programs with their own COVs (such as MRI proposals).*

Manner in which reviewed actions were selected:

Large Awards and Cooperative Agreements (~85 jackets): For large awards (typically about $5,000,000 or more), OAC typically uses the Cooperative Agreement instrument. This instrument allows for greater oversight than the management of a Standard Grant. In some cases, OAC has used the Continuing Grant mechanism, which is a multi-year award in which funds are released in annual increments. Given the size of these awards, OAC has made all of the associated jackets available to the COV.

Standard and Continuing Grants (~235 jackets): Awards and declines (i.e., total competitive actions) have been sampled. Standard practice is to provide at least 40 actions across the entire portfolio of each thematic area (e.g., Data, Software, etc.) for the period covered by the COV. If a given program or thematic area is very small or comparatively recent, all actions have been included in the set. Thus, the resulting set of jackets is a combination of 100 percent coverage or random sampling of actions.

EAGER and RAPID Grants (~ 30 jackets): Despite being a small portion of the actions OAC processes each FY, a sampling of Early-Concept Grants for Exploratory Research (EAGER) and Rapid Response Grant (RAPID) actions are included as an example of the diversity of the portfolio OAC manages. These are small (< $300k) awards to investigators to explore potentially transformative research or to respond to post-disaster rapid research needs. EAGER and RAPID proposals are internally reviewed.

COV Membership

	Name	Affiliation
COV Chair or Co-Chairs:	Juan C. Meza (Chair)	University of California, Merced
COV Members:	Deborah Frincke Rick Arthur Tracy Futhey Grace Wang William Harrod	National Security Agency GE Research Duke University State University of New York Department of Energy

MERIT REVIEW CRITERIA

An understanding of NSF's merit review criteria is important in order to answer some of the questions on the template. Reproduced below is the information provided to proposers in the Grant Proposal Guide about the merit review criteria and the principles associated with them. Also included is a description of some examples of broader impacts, provided by the National Science Board

1. Merit Review Principles
These principles are to be given due diligence by PIs and organizations when preparing proposals and managing projects, by reviewers when reading and evaluating proposals, and by NSF program staff when determining whether or not to recommend proposals for funding and while overseeing awards. Given that NSF is the primary federal agency charged with nurturing and supporting excellence in basic research and education, the following three principles apply:

- All NSF projects should be of the highest quality and have the potential to advance, if not transform, the frontiers of knowledge.

- NSF projects, in the aggregate, should contribute more broadly to achieving societal goals. These broader impacts may be accomplished through the research itself, through activities that are directly related to specific research projects, or through activities that are supported by, but are complementary to, the project. The project activities may be based on previously established and/or innovative methods and approaches, but in either case must be well justified.

- Meaningful assessment and evaluation of NSF funded projects should be based on appropriate metrics, keeping in mind the likely correlation between the effect of broader impacts and the resources provided to implement projects. If the size of the activity is limited, evaluation of that activity in isolation is not likely to be meaningful. Thus, assessing the effectiveness of these activities may best be done at a higher, more aggregated, level than the individual project.

With respect to the third principle, even if assessment of Broader Impacts outcomes for particular projects is done at an aggregated level, PIs are expected to be accountable for carrying out the activities described in the funded project. Thus, individual projects should include clearly stated goals, specific descriptions of the activities that the PI intends to do, and a plan in place to document the outputs of those activities. These three merit review principles provide the basis for the merit review criteria, as well as a context within which the users of the criteria can better understand their intent.

2. Merit Review Criteria
All NSF proposals are evaluated through use of two National Science Board approved merit review criteria. In some instances, however, NSF will employ additional criteria as required to highlight the specific objectives of certain programs and activities.

The two merit review criteria are listed below. Both criteria are to be given full consideration during the review and decision-making processes; each criterion is necessary but neither, by itself, is sufficient. Therefore, proposers must fully address both criteria. (GPG Chapter II.C.2.d.(i) contains additional information for use by proposers in development of the Project Description section of the proposal.) Reviewers are strongly encouraged to review the criteria, including GPG Chapter II.C.2.d.(i), prior to the review of a proposal.

When evaluating NSF proposals, reviewers will be asked to consider what the proposers want to do, why they want to do it, how they plan to do it, how they will know if they succeed, and what benefits

could accrue if the project is successful. These issues apply both to the technical aspects of the proposal and the way in which the project may make broader contributions. To that end, reviewers will be asked to evaluate all proposals against two criteria:

- **Intellectual Merit**: The Intellectual Merit criterion encompasses the potential to advance knowledge; and

- **Broader Impacts**: The Broader Impacts criterion encompasses the potential to benefit society and contribute to the achievement of specific, desired societal outcomes.

The following elements should be considered in the review for both criteria:

1. What is the potential for the proposed activity to:
a. Advance knowledge and understanding within its own field or across different fields (Intellectual Merit); and
b. Benefit society or advance desired societal outcomes (Broader Impacts)?
2. To what extent do the proposed activities suggest and explore creative, original, or potentially transformative concepts?
3. Is the plan for carrying out the proposed activities well-reasoned, well-organized, and based on a sound rationale? Does the plan incorporate a mechanism to assess success?
4. How well qualified is the individual, team, or organization to conduct the proposed activities?
5. Are there adequate resources available to the PI (either at the home organization or through collaborations) to carry out the proposed activities?

3. Examples of Broader Impacts

The National Science Board described some examples of broader impacts of research, beyond the intrinsic importance of advancing knowledge.[2] "These outcomes include (but are not limited to) increased participation of women, persons with disabilities, and underrepresented minorities in science, technology, engineering, and mathematics (STEM); improved STEM education at all levels; increased public scientific literacy and public engagement with science and technology; improved well-being of individuals in society; development of a globally competitive STEM workforce; increased partnerships between academia, industry, and others; increased national security; increased economic competitiveness of the United States; and enhanced infrastructure for research and education. These examples of societally relevant outcomes should not be considered either comprehensive or prescriptive. Investigators may include appropriate outcomes not covered by these examples."

[2] NSB-MR-11-22

INTEGRITY AND EFFICIENCY OF THE PROGRAM'S PROCESSES AND MANAGEMENT

Briefly discuss and provide comments for *each* relevant aspect of the program's review process and management. Comments should be based on a review of proposal actions (awards, declinations, returns without review, and withdrawals) that were *completed within the past four fiscal years*. Provide comments for *each* program being reviewed and for those questions that are relevant to the program(s) under review. Quantitative information may be required for some questions. Constructive comments noting areas in need of improvement are encouraged.

I. Questions about the quality and effectiveness of the program's use of merit review process. Please answer the following questions about the effectiveness of the merit review process and provide comments or concerns in the space below the question**.**

QUALITY AND EFFECTIVENESS OF MERIT REVIEW PROCESS	YES, NO, DATA NOT AVAILABLE, or NOT APPLICABLE
1. Are the review methods (for example, panel, ad hoc, site visits) appropriate? Comments: Review methods were consistent and appropriate to the proposal submissions. Panels were utilized in accordance with agency guidelines and small grants (<$300K) typically did not involve panels but often included consultation among program directors as an internal validation process. For the small to midlevel jackets reviewed, the balancing of panel and ad hoc review appears appropriate and in particular tied to size of the award, with larger awards tending to have more panels rather than virtual reviews. Larger programs also tended to have larger panels. Panel summaries provided excellent comments and feedback that were complemented with review analyses. Where applicable, diary notes also added extra details into the decision process. The PD overview was typically excellent, clearly explaining differences of opinion and decision rationale. There appeared to be a few areas where PDs used their own judgement to recommend against the average of the reviewers, but there was a rationale that could be discerned from the overview, and it is wholly appropriate to expect such instances. **Data Source: EIS/Type of Review Module. <u>Extracted in annual reports</u>**	**YES**
2. Are both merit review criteria addressed a) In individual reviews? b) In panel summaries?	 **YES** **YES** **YES**

c) In Program Officer review analyses? Comments: The review panel commends OAC on consistently using both review criteria across the entire portfolio (a separate discussion specific to the use of the broader impact criteria can be found elsewhere in this report). Both review criteria were addressed in individual reviews, panel summaries and review analyses. Additional review criteria were addressed when appropriate by Program Directors. Individual reviewers consistently rated the proposals along these criteria, although the individual reviewer responses and level of specificity varied greatly, as expected. The individual reviews were of high quality with few outliers. The two criteria were both addressed; acceptance decisions seemed somewhat tipped towards the intellectual merit rather than broader impact. That said the broader impact appeared to be a useful distinguisher among equally meritorious ideas. While reviews were generally thorough, continuing grants tended to have minimal peer review, which the COV felt was justifiable. Small review teams included individuals with appropriate expertise lending gravitas through their insightful and thorough comments. OAC did an excellent job in insuring that both review criteria were addressed in the panel summaries. The panel summaries included appropriate amalgamation of the individual review ratings for the criteria and introduced greater uniformity as to the level and detail across panel review summaries. In all of the Program Director review analysis write-ups that were examined at the small to midlevel size, the analysis was thorough. Program Director review analysis added further depth and detail. The Program Directors' review analyses were consistent with the reviews. **Data Source: Jackets**	

3. Do the individual reviewers giving written reviews provide substantive comments to explain their assessment of the proposals? Comments: Most reviews are of a high quality and substantive. However, there was a significant variation in the quality and length of the reviewer comments. Some reviews are very brief with insufficient details and content. Only a very small number of reviewers had comments that seemed inappropriate and none of the samples examined showed those going out to the PI. While individual review detail and commentary varies greatly among reviewers, in aggregate for a given proposal, there is appropriate detail provided regarding specific strengths, weaknesses, and merit of the proposal. As needed and where appropriate, the internal documentation from individual reviews was edited before it was provided to the PI; the COV found these cases to be appropriate and justified. Overall the individual reviewers did an excellent job in their formal submissions.	**YES**

It is clear that the PD was a moderating influence in some cases (comparing original write up with those returned to the PI). **Data Source: Jackets**	
4. Do the panel summaries provide the rationale for the panel consensus (or reasons consensus was not reached)? Comments: The panel summaries were an excellent resource and did an excellent job of explaining the opinions of the reviewers, as well as the discussion in cases where consensus was not reached. In most cases, the panel summaries were thorough in their description of the overall merits of the proposal and identifying specific strengths and weaknesses. In addition, the panel summaries provided the rationale and sufficient details to justify the panel's recommendations. **Data Source: Jackets**	**YES**
5. Does the documentation in the jacket provide the rationale for the award/decline decision? [Note: Documentation in the jacket usually includes a context statement, individual reviews, panel summary (if applicable), site visit reports (if applicable), program officer review analysis, and staff diary notes.] Comments: The COV commends the Program Directors in their thorough documentation of the process and discourse used to arrive at recommendations for individual proposals. Further correspondence with PIs, diary notes, and PD review analysis were thorough and provided a clear rationale for decisions, including in those cases where panels were not uniform in their reviews or when the PD had further information relevant to the proposal. It was easy to determine the rationale for the award/decline decision. There were only a few decline decisions slightly lacking in justification. **Data Source: Jackets**	**YES**

6. Does the documentation to the PI provide the rationale for the award/decline decision? [Note: Documentation to PI usually includes context statement, individual reviews, panel summary (if applicable), site visit reports (if applicable), and, if not otherwise provided in the panel summary, an explanation from the program officer (written in the PO Comments field or emailed with a copy in the jacket, or telephoned with a diary note in the jacket) of the basis for a declination.] Comments: PIs were typically provided with individual reviews as well as the panel summary, each noting strengths, weaknesses, merit and broader impact. These generally provide sufficient rationale in cases where award/decline decisions aligned closely along the spectrum of High Competitive-Competitive-Low Competitive-Non Competitive. Overall, the documentation was well done and the material provided to the PI was largely of high quality. The documentation that is sent to the PI provides rationale that is consistent with the comments made by the reviewers. Often comments are provided that would help the proposer develop a stronger proposal for a future solicitation. The COV did note that some review analyses did not have clear Context Statements. **Data Source: Jackets**	**YES**
7. Additional comments on the quality and effectiveness of the program's use of merit review process: The COV was impressed by the quality and effectiveness of the merit review process. OAC Program Directors are diligent and thoughtful in their application of this process and in ensuring appropriate documentation of the process and rigorous expert and community review of proposal. It is clear that intellectual merit was the strongest driver of the decisions. In some cases, decisions between similarly ranked proposals at the small/midlevel tipped based on the broader impact. The OAC should recognize the reviewer bench of experience as a significant asset and maintain these relationships.	

II. Questions concerning the selection of reviewers. Please answer the following questions about the selection of reviewers and provide comments or concerns in the space below the question.

SELECTION OF REVIEWERS	YES, NO, DATA NOT AVAILABLE, or NOT APPLICABLE
1. Did the program make use of reviewers having appropriate expertise and/or qualifications? Comments: The Program Directors did an excellent job selecting reviewers with the required technical expertise. Reviewers had appropriate expertise/qualifications for the proposal evaluations they conducted. Some had extensive operational experience. Others were more research focused in their areas of expertise. The COV notes that for some programs and specific proposals, the available pool of reviewers with the necessary expertise can be quite limited. We commend the Program Directors in their success in getting an appropriate pool of experts, but note that the diversity of the pool often suffered as a result, with many panels lacking gender and racial diversity. In at least some areas there appeared to be a preference to ensure that the industry-lab proposals included at least a few industry-lab reviewers. However, the participation of industry experts in panels seem to be low particularly considering the mission of OAC. It would be preferable to have more industry-lab reviewers throughout but this might be unrealistic given the reviewer pool. In the cyberinfrastructure area, it is critical to understand the state of the industry art, and it may not be possible to get there solely by bringing in industry scientists. This is an area for ongoing consideration regarding causes of reluctance, availability, and/or relationships with or knowledge of appropriate individuals for potential reviewers. Likewise, gender balance in reviewers was not strong in particular areas. Although this did not have an adverse impact on the outcomes, it might negatively bias future competitiveness. The level of participation of underrepresented minorities, was more difficult to assess as the available data is more difficult to gather. **Data Source: Jackets**	**YES**
2. Did the program recognize and resolve conflicts of interest when appropriate? Comments:	**YES**

In those occasional cases where conflicts of interest were identified or arose, they were appropriately resolved and documented within the review analysis summaries. Several cases of potential conflict of interest were discussed in the jackets. In all instances examined, the conflict was caught relatively early, appropriately adjudicated, and noted in a clear way. Relevant information was captured in the summary, the diary, and the conflicts section. Reviewer conflicts were well documented and appropriate actions were taken to resolve the conflict. **Data Source: Jackets**	
3. Additional comments on reviewer selection: As noted above, the diversity of the review panels often lacked gender and/or racial diversity. The well-documented dearth of diversity in STEM fields notwithstanding, we urge NSF leadership to consider whether alternative methods might be employed to address this issue, such as through the use of pre-proposals, triage or other automated means of winnowing the set of full proposals that need to be empaneled, thereby reducing the demand for such a large list of reviewers. Increasing the reviewers could also be a means to systematically grow hands-on experience and confidence among more junior professionals as well as include a broader diversity while maintaining confidence in the overall thoroughness of reviews. It may make sense to increase the number of virtual reviews or even to hold reviews in person at alternate sites. In particular, rotating panel locations near industry centers in software and hardware might encourage participation by panelists from industry, broadening the knowledge of the overall panel to be more comprehensive, and specifically with regards to applied experience. Alternately, proposals at a high-dollar level might benefit by a follow up review by specific industry participants. This may be occurring but was not noted in any of the reviewed jackets. It would have been great if OAC had provided the overall statistics of the number of unique reviewers, the number of panels and the number of reviews, and the institutional distribution of reviewers in the last three years. The anecdotal observation is that there appears to be a reasonable gender diversity among reviewers, but lacking of industry participation and URMs.	**YES**

III. Questions concerning the management of the program under review. Please comment on the following:

MANAGEMENT OF THE PROGRAM UNDER REVIEW
1. Management of the program. Comments: The overall management of the program appears to be excellent. The entire office is highly responsive to the rapidly changing environment and appears to make diligent efforts to both learn what the community needs are, and to plan and prioritize the activities within OAC. Overall, the OAC programs were very well managed. The due diligence and the quality of leadership team and program directors are impressive. The OAC could benefit from even more strategic thinking when formulating solicitations and DCLs, although the COV understands that this will require more time and resources than OAC may be able to afford. For the equipment acquisitions programs, there were four projects funded. The contract value ranges from $6M to $30M. These are clearly important projects that require a higher level of project management. Routine calls among participants were consistently and appropriately utilized (weekly/by-weekly/monthly, where appropriate) for ongoing facility or equipment grants, and site visits provided further validation.
2. Responsiveness of the program to emerging research and education opportunities. Comments: The program's responsiveness to emerging research and education opportunities is excellent. The Office uses a variety of strategies to learn about the needs both within the Foundation and the community at large. Examples of these types of activities include participation in NSF cross-directorate panels and working groups (HDR, Quantum Learning, Arctic) as well as external groups including NITRD and participation in interagency reviews. In addition, the Program Directors attend annual PI meetings and professional conferences in their fields to stay abreast of the latest trends within their fields. In this context, the OAC and NSF should continue to support the utilization of IPAs as a hiring mechanism; continue to conduct Annual PI meetings; and expand PD trainings as well topical workshops to ensure that OAC scientific team stay atop the most recent research, technological and education trends and needs.
3. Program planning and prioritization process (internal and external) that guided the development of the portfolio. Comments:

OAC has a well-defined set of activities to help it plan and prioritize its various programs. In addition to the ones listed above, the Office also has regular meetings with its advisory committee (ACCI), which provides external advice on the programs as well as recommendations from various working groups on topics related to cyberinfrastructure.

GW - The program planning and prioritization process exist, but not as well defined.

4. Responsiveness of program to previous COV comments and recommendations.

Comments:

The previous COV provided several findings and recommendations with regard to the dwell time for some of the larger grants, the size and duration of the DataNet awards, the tracking of outcomes of the Taskforce reports from 2009, the skill sets required for the management of the large-scale programs such as Track 1 and DataNet, and the turnover in some of the program staff. In each case, OAC provided an appropriate and thoughtful response to the findings/recommendations.

The OAC has been generally responsive to the DataNet concern raised by the last COV. The OAC leadership team and Program Directors are encouraged to consider lessons learned from the DataNet experience to further enhance the transparency and communication with the community and formulate a strategy to establish proper award size and duration by taking the budget constraints and community needs into consideration.

As many of the findings reflect both the rapidly changing environment (both technically and fiscally) and in some cases longer-term projects, we encourage the Office to continue to monitor and address the issues highlighted in earlier COV reports.

The COV also noted that the reviewer reports and metadata available to this COV seemed noticeably improved in more recent years (2016 in particular).

IV. Questions about Portfolio. Please answer the following about the portfolio of awards made by the program under review.

RESULTING PORTFOLIO OF AWARDS	APPROPRIATE, NOT APPROPRIATE, OR DATA NOT AVAILABLE
1. Does the program portfolio have an appropriate balance of awards across disciplines and sub-disciplines of the activity? Comments: The program portfolio covers a wide spectrum of activities that align with OAC's mission and appears to be well balanced across disciplines. In areas such as International Networking, there has been conscious consideration of the need to support a wide range of science domains and geographies. The learning and workforce development component is small and it needs to be strategically determined what OAC's focus should be in this area. The NSF might consider a survey report assessing computational maturity and engagement of cyberinfrastructure across scientific disciplines to identify opportunities for outreach or to capture best practices that would stimulate adoption of cyberinfrastructure over a broad range of scientific disciplines. **Data Source: EIS/Committee of Visitors Module. From the Report View drop-down, select the Funding Rate module to see counts of proposals and awards for programs. The Proposal Count by Type Report View will also provide a summary of proposals by program. Extracted in annual reports**	**YES**
2. Are awards appropriate in size and duration for the scope of the projects? Comments: The awards in each category viewed appeared to be appropriate given the scope. In addition, the size and duration of awards was well correlated with the scope of the projects and the complexity and constraints related to hardware acquisitions, network deployments and/or software development. In most cases observed, the appropriateness of the award size was noted in the reviewer summary, or questioned in some cases in the diary dialog with the PI. The COV did find that review analyses did not consistently address whether the budget and award durations were appropriate. The degree to which reviewers are expected to opine on pass/fail vs. reward sizing or suggesting conditional remarks (e.g., an E rating at reduced funding versus a G rating at proposed funding) could be clarified to COV and reviewers.	**YES**

Data Source: EIS/Committee of Visitors Module. From the Report View drop-down, select Average Award Size and Duration. <u>Extracted in annual reports</u>	
3. Does the program portfolio include awards for projects that are innovative or potentially transformative? Comments: The program portfolio includes numerous examples of innovative and/or transformative projects, but the term (3 years) the COV was asked to review limited the ability for a thorough longitudinal reflection. It might be worthwhile to do consider a retrospective (10 years) as to how innovative and transformative awards were. **Data Source: Jackets**	**YES**
4. Does the program portfolio include inter- and multi-disciplinary projects? Comments: The COV was impressed by the extent to which OAC has been successful in co-funding programs with other NSF directorates. There are were several inter/multi-disciplinary programs noted in the PD briefings and in the annual reports. This is an appropriate focus of the program and is one of the areas where the Science Advisor position has been helpful. Overall, the program portfolio covers a broad spectrum of inter- and multi-disciplinary projects that have the potential to transform research in many areas. The term trans-disciplinary has recently received increasing attention. In the context of OAC providing the means by which this distinction may be facilitated – perhaps OAC should consider its implications in a continuum relative to inter and multi-disciplinary. **Data Source: If co-funding is a desired proxy for measuring inter- and multi-disciplinary projects, the Co-Funding from Contributing Orgs and Co-Funding Contributed to Recipient Orgs reports can be obtained using the EIS/Committee of Visitors Module. They are available as selections on the Report View drop-down. <u>Extracted in annual reports</u>**	**YES**
5. Does the program portfolio have an appropriate geographical distribution of Principal Investigators? Comments: This area was hard to analyze with the data available. Overall, the geographical distribution was fairly broad and any variation was consistent with the COV's experience and expectations. A fuller analysis that includes data that captures individual institutions might prove useful.	**YES**

Data Source: EIS/Committee of Visitors Module. Select Proposals by State from the Report View drop-down. Extracted in annual reports	
6. Does the program portfolio have an appropriate balance of awards to different types of institutions? Comments: In general, awards are balanced to different institutional characteristics; however, we do note that research universities are disproportionately successful as it relates to large grant proposals. The COV does not find this to be problematic in and of itself, but notes it is an area of imbalance. The portfolio contains a large share of R1 institutions, with fewer non-R1 institutions. The largest grants tend to go to well-known institutions (fully supported by the reviews) with well-known PIs. The annual report for 2016 shows that the preponderance of grants went to Top 100 research intensive PhD institutions, which also write more than half the total proposals. **Data Source: EIS/Committee of Visitors Module. Select Proposals by Institution Type from the Report View drop-down. Also, the Obligations by Institution Type will provide information on the funding to institutions by type. Extracted in annual reports**	**YES**
7. Does the program portfolio have an appropriate balance of awards to new and early-career investigators? NOTE: A new investigator is an individual who has not served as the PI or Co-PI on any award from NSF (with the exception of doctoral dissertation awards, graduate or post-doctoral fellowships, research planning grants, or conferences, symposia and workshop grants.) An early-career investigator is defined as someone within seven years of receiving his or her last degree at the time of the award. Comments: The portfolio seems to have an appropriate balance between new and early-career investigators and more senior PIs. Based on the data provided the funding rates for new versus prior PIs appears to be similar and the only major difference is in the number of proposals submitted. In conversations with OAC during the review, this seems to be an area that they pay particular attention to. The OAC leadership also discussed with the COV the challenges in their ability to collect the relevant demographic data, as some of the data is self-reported as well as voluntary. OAC is encouraged to continue to cultivate research leadership among early-career investigators. **Data Source: EIS/Committee of Visitors Module. Select Funding Rate from the Report View drop-down. After this report is run, use the Category Filter button to select New PI for the PI Status filter or New Involvement (PIs & coPIs) = Yes. Extracted in annual reports**	**YES**

8. Does the program portfolio include projects that integrate research and education? Comments: There typically was a discussion of broader impacts of each proposal, as well as REU and Early Career grants, and a very small number of education grants. The education-specific (as purpose of the grant) investment seems quite small. Although, the education component was included in many projects it does not represent a strong component overall. However, the COV noted that there is a newer program looking at Workforce Development, intended to advance the workforce that explicitly requires proposals to support training/education goals. This is a critical area and it is laudable to have it within OAC. **Data Source: Jackets**	**YES**
9. Does the program portfolio have appropriate participation of underrepresented groups[3]? Comments: Women and minorities continue to be relatively underrepresented in STEM, based on internal documents made available to the COV. Among reporting females and minorities, the success rate tended to be slightly better for minorities than non-minorities, and significantly better for females. Likewise, disability demographics showed few reporting a disability; but we did note a slightly improved acceptance rate. There is clearly an issue; but the issue seems to be within the pipeline and not how the proposals are handled once received. **Data Source: EIS/Committee of Visitors Module. Select Funding Rate from the Report View drop-down. After this report is run, use the Category Filter button to select Women Involvement = Yes or Minority Involvement = Yes to apply the appropriate filters. <u>Extracted in annual reports</u>**	**YES**
10. Is the program relevant to national priorities, agency mission, relevant fields and other constituent needs? Include citations of relevant external reports. Comments: Numerous reports both within and outside NSF point to the need for a robust and state-of-the-art cyberinfrastructure to support the research needs of science and engineering. The most recent such report by the National	**YES**

[3] NSF does not have the legal authority to require principal investigators or reviewers to provide demographic data. Since provision of such data is voluntary, the demographic data available are incomplete. This may make it difficult to answer this question for small programs. However, experience suggests that even with the limited data available, COVs are able to provide a meaningful response to this question for most programs.

Academies, Future Directions for NSF Advanced Computing Infrastructure to Support U.S. Science and Engineering in 2017-2020 makes an excellent case for this. In that report, their first recommendation states, *"The National Science Foundation (NSF) should sustain and seek to grow its investments in advanced computing—to include hardware and services, software and algorithms, and expertise—**to ensure that the nation's researchers can continue to work at frontiers of science and engineering**".* It is difficult to imagine a situation in which the U.S. could maintain its science leadership role without the type of investments that OAC makes and the role that is plays in supporting an advanced cyberinfrastructure. **Data Source: Jackets**	
11. Additional comments on the quality of the projects or the balance of the portfolio: This is a high quality program supporting critical scientific discovery, and likewise increasing access to the CI required to perform high-end scientific discovery in many areas.	

OTHER TOPICS

1. Please comment on any program areas in need of improvement or gaps (if any) within program areas.

- As in the past, the technologies that underlie advanced cyberinfrastructure continue to evolve rapidly, presenting a persistent challenge to deliver a stable cyberinfrastructure despite such volatility. Propelled by the greater abundance and affordability of computing, storage and network performance, a wider range of applications have embraced computational methods and scalable data analytics. Among the emerging communities shifting use away from the established incumbent users, the application of deep learning shows strong momentum. Responding to this trend, commercial vendors of both hardware and hosting elastic environments (cloud) have significantly advanced the ability of their platforms to deliver high performance training for machine learning. The potential for machine perception and data analysis to broadly transform the practice of science itself is evident, placing OAC at the heart of this revolution.

    - Thus, the OAC mission faces challenges in:
    - Maintaining expertise in advanced computing hardware
    - Sustaining software compatibility and scalability to fully leverage state of the art hardware
    - Advising both incumbent and newly-interested user communities in readiness of maturing technologies for desired applications as well as potential to overcome prior limitations

- OAC is not alone in these challenges, however, and could significantly benefit from alignment with other NSCI participants. Further, the commercial elastic infrastructure (cloud) in 2018 may offer hardware and software architecture significantly closer to NSF mission needs than in prior reviews. For example, one major cloud provider's recent partnership with a supercomputing vendor to co-locate leadership-class architecture with elastic compute and storage.

- There appears to be some inconsistencies in how the data management plans are reviewed, discussed, and managed across the program areas.

2. Please provide comments as appropriate on the program's performance in meeting program-specific goals and objectives that are not covered by the above questions.

- It might be useful to develop a target metric to assess the access to necessary scientific equipment by target institution and discipline, and report this to future COVs. This data would help indicate progress in ensuring diverse participation in the scientific process.

3. Please identify agency-wide issues that should be addressed by NSF to help improve the program's performance.

- Better automated tools for program workflow, review, metadata capture, standardization, management, assessment and monitoring.
- Augmented perception / augmented intelligence as a technique itself or in application to other contexts
- Are there processes considering how OAC may support AI-as-a-collaborator or AI-as-assistant, similar to efforts such as Cortana or other domain-specific machine expert systems.

- To what extent can tool-enabled collaboration promote cross-disciplinary to multi-disciplinary to trans-disciplinary (holistic) teaming?
- Cyberinfrastructure transforming the practice of science utilizing (assiduous, patient, consistent, N-dimensional, …) beyond-human capabilities of computational systems.

4. Please provide comments on any other issues the COV feels are relevant.

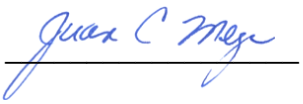
- It might prove useful to investigate what post-mortem information could be captured throughout execution of program creation, reviews, execution and completion tasks could be captured for ongoing inspection and improvement of the NSF processes? (in particular leveraging a knowledgebase for institutional memory on practices that worked or concerning interactions with individuals or institutions – like a CRM).
- NSF might consider benchmarking the introduction of information systems within the FDA in improving review of certification of medical devices.

5. NSF would appreciate your comments on how to improve the COV review process, format and report template.

- It would be useful to have statistics across entire portfolio in addition to the selection of jackets provided.
- COV could be much more efficient, responsive, and accurate if it had better summaries of data for easier analyses. (See appendices for some examples.)
- A list summary of awards (MyCOV) would be useful in spotting patterns of interest or in discovery of good drill-down targets by providing the following – in an Excel form if not in MyCOV itself, with an option for a summary of each time that brings everything together in a single view/click (See Appendix A.)
- Documents should all be in commonly available formats such as HTML or PDF. For example, some eJacket documents may not be openable by a reviewer using their own laptop if they lack Microsoft Office or the laptop has security installed that prevents opening of certain files.
- COV recommends that OAC/NSF develop more robust systems for maintaining a historical record of the data and documentation provided to COVs as well as supplementary materials to better evaluate longitudinal trends.

The Committee of Visitors is part of a Federal advisory committee. The function of Federal advisory committees is advisory only. Any opinions, findings, conclusions, or recommendations expressed in this material are those of the Advisory Committee, and do not necessarily reflect the views of the National Science Foundation.

SIGNATURE BLOCK:



For the Office of Advanced Cyberinfrastructure Committee of Visitors
Juan C. Meza
Chair

Appendix A.

With respect to additional information for future COVs, a list summary of awards (MyCOV) in spotting patterns of interest would prove useful in discovery of good drill-down targets by providing the following – in an Excel form if not in MyCOV itself:

Id	PI Name	PI Institution	PI Wins	PI Losses	Award Program	Weeks Receipt to Review	Weeks Receipt to Award/Decline	# of Reviewers	# of Correspondences	Requested $	Awarded $ (or 0)	Reviewer Ratings	Final Ratings
1255826	Carset, Thomas M	UNC Chapel Hill	8	17	12-557	24	30	4	0	109000	109000	VG G VG E	E VG

This could be followed by a summary of each entry that brings together in a single view (click):

PI		Name	Institution		Awards	Declines			
Timeline		Submit	Review	Response					
Funding		Requested	Awarded						
Reviewers									
		Name	Institution	Doc link	C1-Rating	C2-Rating	C3-Rating	C4-Rating	C5-Rating
		Name	Institution	Doc link	C1-Rating	C2-Rating	C3-Rating	C4-Rating	C5-Rating
		Name	Institution	Doc link	C1-Rating	C2-Rating	C3-Rating	C4-Rating	C5-Rating
Final				Doc link	C1-Rating	C2-Rating	C3-Rating	C4-Rating	C5-Rating

Learning opportunities in the NSF processes

For example: an internal knowledge base reviewers and administrators can leverage for:

- Methods to assess success & capturing success metrics
    - Not simply capturing the “benefits of success” but the “benefits of failure”.
    - Opportunity to capture “years-out” impact such as dependent work, financial impacts, growth in jobs/students, etc.
- Availability of validation data & characterization into a common listing to facilitate cross-referencing or use in future training sets or control sets (note ASME has working groups on virtual validation)
- Better standardizing expected state of practice for:
    - Data conventions, archival, integrity, formatting, protection, etc.
    - Software libraries, license terms, build/test/deployment automation, containerization, workflow, etc.
- Means to facilitate reproducibility (relating to both the data and software aspects) Application of Automation Tools – including application of Machine Learning / Deep Learning (which has blossomed since the timeframe of the prior COV assessment.)

OAC mission: Assess how human + cyberinfrastructure collaboration may be fundamentally altering the overall NSF mission

- What are critical barriers to the NSF mission that can be reduced by leveraging automation? (previously impractical or tedious tasks that could now be performed via softbots?)