



National Science Foundation
4201 Wilson Boulevard
Arlington, Virginia 22230

NSF 12-090

Dear Colleague Letter: New Solutions to Create Integrative Data Management Infrastructure(s) for Research Across the Sciences

NSF seeks transformative approaches to create integrative data management infrastructures across science and engineering disciplines that will advance US research and innovation through new capabilities empowering the majority of the nation's researchers whether they are working individually, in small groups, or as part of large projects.

These efforts will enhance our understanding of the scope and nature of legacy, current, and potential future data sets in Science, Technology, Engineering, and Mathematics (STEM). The opportunity has never been greater, and NSF seeks input from the community.

NSF places significant emphasis on computational and data-rich science and engineering, and recognizes the complex challenges in many such disciplines, for example, that the volume and fragmented nature of generated data sets is often difficult to handle with traditional tools. Transformative approaches and innovative technologies are needed for heterogeneous data to be integrated, made interoperable, explored and re-purposed by researchers in disparate fields and for myriad uses across institutional, disciplinary, spatial and temporal boundaries.

NSF offices and directorates make substantial investments in producing and collecting data through NSF-supported research facilities and projects, as well as helping the science and engineering community utilize data collected by other entities around the world. Similarly, OCI makes substantial investments in advanced high-performance computing, data infrastructure, software development, virtual organizations and networking. The present partnership among these divisions seeks to integrate these data and technologies in open, adaptable, and sustainable integrative framework(s) that will support transformative research and foster the development of new data-focused methodologies.

New tools will be needed to manage a collaborative effort and build community support. We seek to find commonalities in the data needs across the disciplines and to develop effective approaches to address these needs. It is expected that collaborative efforts will emerge among researchers across different disciplines in response to these challenges. The solutions must constructively provide opportunities for integrating the large numbers of smaller projects as well as engage major NSF-funded research and infrastructure projects. The effort should also be coordinated with other Federal and international agencies and foster symbiotic relationships that will be essential for NSF to consider investment in a prototype enterprise.

NSF will host a series of webinars followed by a charrette meeting to rapidly facilitate the early stages of novel approaches to achieve the goals described. NSF expects that actionable ideas emerging from the charrette will help define future developments of this system in FY2013 and beyond. The webinars will begin on June 15, 2012 (details on how to register for participation in the June 15, 2012 webinar are given below). This dialog with the community will provide context for the initial integrative data-focused infrastructure vision and offer an opportunity for the community to interact with the leadership of the NSF units involved. Additional details on the sequence of events will be provided at the webinar and will be posted on the MPS, BIO, CISE, EHR, ENG, GEO, OCI, and SBE websites.

Signed

Edward Seidel
Assistant Director
Directorate for Mathematical and Physical
Sciences

Alan Blatecky
Office Head
Office for Cyberinfrastructure

Joan Ferrini-Mundy
Assistant Director
Directorate for Education and Human
Resources

Myron Gutman
Assistant Director
Directorate for Social, Behavioral And Economic
Sciences

Farnam Jahanian
Assistant Director
Directorate for Computer and Information
Science
and Engineering

Tim Killeen
Assistant Director
Directorate for Geosciences

Thomas Peterson
Assistant Director
Directorate for Engineering

John Wingfield
Assistant Director
Directorate for Biological Sciences

INVITATION FOR JUNE 15, 1PM-3PM EASTERN TIME - DEAR COLLEAGUE LETTER VIRTUAL MEETING

We invite you to attend this online meeting.

Topic: DataWay- Dear Colleague Letter Virtual Meeting
Friday, June 15, 2012

Date: 1:00 pm, Eastern Daylight Time (New York, GMT-
Time: 01:00)

To register for this meeting

1. To register, go to
<https://mmancusa.webex.com/mmancusa/j.php?ED=183244797&RG=1&UID=1279493797&RT=MiMxMQ%3D%3D>
2. Enter required registration information (only fields marked with asterisks are required).

Once the host approves your request, you will receive a confirmation email with instructions for joining the meeting. The instructions will include a URL to view real-time video and slide presentations over the web and a telephone number with participant pass code to connect to the audio bridge to participate in real-time audio discussions. Upon joining the audio bridge, participants will be placed in listen only mode during the main presentation. A question and answer session will follow the main presentation. Participants will be instructed how to enter the queue to ask questions.

To view in other time zones or languages, please click the link:

<https://mmancusa.webex.com/mmancusa/j.php?ED=183244797&RG=1&UID=1279493797&ORT=MiMxMQ%3D%3D>

For assistance

1. Go to <https://mmancusa.webex.com/mmancusa/mc>
2. On the left navigation bar, click "Support".

Questions about the Dear Colleague Letter and the virtual meeting can be sent to DataWay@nsf.gov. You also can contact Paul Spyropoulos at: pspyropo@nsf.gov or 1-703-292-4968

The playback of UCF (Universal Communications Format) rich media files requires appropriate players. To view this type of rich media files in the meeting, please check whether you have the players installed on your computer by going to <https://mmancusa.webex.com/mmancusa/systemdiagnosis.php>

<http://www.webex.com>

IMPORTANT NOTICE: This WebEx service includes a feature that allows audio and any documents and other materials exchanged or viewed during the session to be recorded. By joining this session, you automatically consent to such recordings. If you do not consent to the recording, do not join the session.