The NSF Principal Investigator meeting brings together project participants that are funded through the same National Science Foundation program in order to foster interactions. The Coming Together Around Data workshop brought together teams funded from two NSF programs, the Digital Data Preservation and Access Network Partners (DataNet) program and the Community-based Data Interoperability Networks (INTEROP) program. In an area that is moving as rapidly as is the interoperability, sharing, and archival of scientific data, a PI meeting is a venue for accomplishing more and varied kinds of objectives. We hosted an unconventional PI meeting/workshop with a two-fold goal: to foster exchange of information about projects in the DataNet and INTEROP programs, and 2.) to plant first seeds for long lived, collective effort that advances scientific sharing and preservation worldwide.

Nearly 60 participants attended the meeting, shown in the included picture taken during the workshop. Participants included four NSF officers, representatives from NOAA, NASA, Germany, Netherlands, the Library of Congress, 10 DataNet programs and 9 INTEROP programs. It also included Ph.D. and Master's students and postdoctoral fellows largely from Indiana University who helped with the workshop and in process learned from experts in the field of data science.

The intellectual contributions the workshop were achieved through a program that drew on notions from the "unconference" where topics emerge spontaneously with the goal of fostering emergent ideas or efforts. A number of topics emerged as important for discussion. On day one of the workshop, five separate subgroups were formed on the following topics:

- Storage and funding models; clouds, libraries, repositories
- Metadata and semantics
- Building a lasting coordination group (Science Data Consortium)
- Data citation
- Education and curriculum for data curation

On the second day new topics emerged while some of the topics emerged a second time for continued discussion. These topics included

- Science Data Consortium
- Storage and funding models; clouds, libraries, repositories, cont.
- What will Data Realm look like in 20 years?
- Interoperability between DataNet infrastructures: what is possible today?

The notion of the Science Data Consortium first emerged from the workshop with the following guiding principles: It must be open in that all can contribute effort, ideas, code. The deliverables or objectives of a Science Data
Consortium are to be knowledge exchange, and recommendations/standards (either with or without supporting code) for use to simplify work of domain scientists. Information is to be delivered to the communities we serve. The Charter is an early deliverable. The interested participants from the workshop preferred a stronger grassroots orientation, at to start. It was decided as a next step to establish a small group to clarify the vision and move things forward.

The advisory group once established post workshop went on to undertake a number of outreach activities, described below, that eventually contributed to the Research Data Alliance http://rd-alliance.org/.


- Birds of a Feather (BoF) on Data Set Identifier Interoperability held at Internet Engineering TaskForce (IETF) 84 in Vancouver, BC, Canada, July 29 – August 3, 2012.

NSF funding through the DataNet and INTEROP programs has created and brought together a powerful element of technology contributors and researchers who deeply understand and are trying to advance the BigScience scientific data management worldwide. The collective effort of the workshop allowed synergies, best practices, and standards to emerge. As witnessed by the abundant quality of follow-on activities which came out of the meeting, the time is right.