

**DOE EPSCOR
Award Number DE-SC0004851
Final Report
Report Number DE-SC0004851-IIa-1**

**Exploiting the Use of Social Networking to
Facilitate Collaboration in the
Scientific Community**

Principal Investigator – Edrick G. Coppock

Date: April 7, 2014

**Period of Report
7/15/2010 through 7/14/2014**

Total Approved Budget - \$554,769.00

**Collaborating Institution: Oak Ridge National Laboratory (ORNL) Computational
Sciences and Engineering Division: PI- Mallikarjun Shankar, Ph.D.**

Submitted By

Information International Associates, Inc.
104 Union Valley Road
P.O. Box 4219
Oak Ridge, TN 37831-4219
865-481-0388 Voice
865-481-0390 Fax
www.iiaweb.com

Section 5. Abstract of project goal and objective.

The goal of this project was to exploit social networking to facilitate scientific collaboration. The project objective was to research and identify scientific collaboration styles that are best served by social networking applications and to model the most effective social networking applications to substantiate how social networking can support scientific collaboration. To achieve this goal and objective, the project was to develop an understanding of the types of collaborations conducted by scientific researchers, through classification, data analysis and identification of unique collaboration requirements. Another technical objective in support of this goal was to understand the current state of technology in collaboration tools. In order to test hypotheses about which social networking applications effectively support scientific collaboration the project was to create a prototype scientific collaboration system. The ultimate goal for testing the hypotheses and research of the project was to refine the prototype into a functional application that could effectively facilitate and grow collaboration within the U.S. Department of Energy (DOE) research community.

In earlier progress reports, we reported, with the introduction and availability of several collaboration tools that are well-endowed with venture capital, that it would not be possible to develop a competitive collaboration tool given the amount of funding and time remaining on this grant. We also reported that we had developed a survey instrument that would provide insight into collaboration styles and the degree of adaption of social media in scientific collaboration. Unfortunately, DOE management determined that the exercise of the instrument would violate the Paperwork Reduction Act and, thus, we aborted that part of the research.

As a result, we concentrated on the following objectives for the remainder of the grant performance period:

1. Research the scientific collaboration within the DOE research community using DOE data from the DOE research databases.
2. Perform cleansing and disambiguation of the DOE data.
3. Using the cleansed data, perform a statistical analysis of collaboration habits and styles by examining a large number of DOE researchers and look at collaboration across geographies, subject matter and timelines to gain a better understanding of those collaboration habits.
4. Develop collaborative relationships with other purveyors of scientific research in order to capture and understand their perspective on the role of social media in scientific research.
5. Produce approximately five papers detailing the results of the research above to be submitted to peer-reviewed journals for publication.
6. Write the final report and submit to DOE Energy Link System (E-Link) with the appropriate DOE Form 241.3.

Description of Accomplishments

In Year 1, the team focused on 3 research areas:

1. Understand Collaboration Types/Styles

2. Assess current state of technology in collaboration tools
3. Identify requirements for a social networking collaboration prototype

In the Year 1 Progress report, we provided the details of that research and the findings of the team.

In Year 2, our efforts were focused to:

1. Research open source technologies and other existing technologies as the basis for creating a prototype professional profile framework.
2. Create initial prototype as a model for testing collaboration features and for feedback.

After evaluating the requirements and seeing the introduction of several new collaboration tools with many more under development, the team determined that we should concentrate our research on better understanding of collaboration types rather than on building another collaboration tool. To support this research, we developed a survey instrument to learn more about collaboration habits of scientists. However, we were unable to exercise the instrument due to the Department of Energy's decision that the survey would not be in compliance with the Paperwork Reduction Act.

In Year 3, the team decided to spend the remaining time focusing on collaboration trends in DOE and across other Federal and Commercial organizations. These efforts are summarized in the following paragraphs.

1. Research the scientific collaboration within the DOE research community using DOE data from the DOE research databases.

In order to discover meaningful information on the collaboration trends within the DOE research community, the team decided to apply powerful statistical research methodology to the vast storehouses of DOE research data. The data selected for the research was from the DOE Information Bridge and included more than 1.2 million records.

2. Perform cleansing and disambiguation of the DOE data.

Over the past several years, DOE has moved from entering a number of data fields with a controlled thesaurus or authority file to freeform. Owing to the power of modern search engines, this does not significantly detract from the user search experience. However, it does create certain ambiguities in several fields. In order to create more meaningful statistical results, a good deal of effort was required to cleanse and disambiguate the data. This work was primarily performed with graduate student labor.

3. Using the cleansed data, perform a statistical analysis of collaboration habits and styles by examining a large number of DOE researchers and look at collaboration across geographies, subject matter and timelines to gain a better understanding of those collaboration habits.

This work was performed by Professor Hampersum and his graduate/postgraduate staff. The results were collected and documented in a paper titled "*Social Network Analysis of Scientific Collaborations across Different Subject Fields*" which was published in an international journal that is described below.

4. Develop collaborative relationships with other purveyors of scientific research in order to capture and understand their perspective on the role of social media in scientific research.

For this part of the research, we reached out to the following organizations with the activities described:

- U. S. Department of Defense Technical Information Center (DTIC) - We obtained an account on their flagship collaboration tool, Aristotle, evaluated the capabilities and compared them with other collaboration tools. In addition, we discovered that DTIC used social media in a number of ways to promote scientific collaboration. As a result, DTIC agreed to provide a paper describing their collaboration tools and experiences for submittal to the international journal described below.
- PeerJ Inc. - PeerJ Inc. is the Open Access publisher of *PeerJ* (a peer-reviewed, Open Access journal) and *PeerJ PrePrints* (an un-peer-reviewed preprint server), both serving the biological, medical and health sciences. In our discussions with PeerJ we believe they have a unique perspective on Open Access publication coupled with novel concepts for scientific collaboration. PeerJ has provided a paper describing their Open Access experience as it applies to scientific collaboration which was submitted to the international journal described below.
- Oak Ridge National Laboratory (ORNL) - ORNL is an uncompensated collaborator on this EPSCoR grant. We have had several discussions about future partnerships around their technology transfer program that look promising to leverage more research funds into this geographic area. In addition, ORNL agreed to provide a paper discussing their experience in scientific collaboration in their Big Data initiative.
- University of Tennessee - Knoxville (UTK) - UTK is a subcontractor on this EPSCoR grant. Exploring research opportunities with the Department of Statistics, Operations, and Management Science and the Department of Mathematics, we have identified and submitted proposals for four research opportunities. We won one of the grants - ***Weather/Climate Variability Impact on Energy, Water and Food Resources and Implications for Regional Stability, US Army SBIR Phase I***, Topic Number OSD09-HS1, Contract Number W913E5-10-C-0012. We are actively pursuing the identification of additional opportunities. In addition, UT has conducted research on scientific collaboration in the DOE research community and has prepared the paper described below.

5. Produce approximately five papers detailing the results of the research above to be submitted to peer-reviewed journals for publication.

In order for this research to gain exposure to the widest audience possible, we sought a Journal that would be an appropriate vehicle to publish this research. As a result, we had discussions with the Information Services & Use Journal (ISU), resulting in agreement to publish a special edition around the concept of the role of social media in scientific collaboration. Information Services & Use is an information and information technology oriented publication with a wide scope of subject matters. International in terms of both audience and authorship, the journal aims at leaders in information management and applications in an attempt to keep them fully informed of fast-moving developments in the information field. To this end, we worked internally as well as with our collaborators mentioned above to produce the following papers that were published in the fall issue of ISU.

- *Status of the Adoption of Social Media in the Scientific Research Community*, Coppock, Edrick G. Information International Associates, Inc., 104 Union Valley Road, Oak Ridge, TN 37831-4219 and Davis, Lynn, U. S. Department of Energy, Office of Scientific and Technical Information, 1 Science.gov Way, P.O.

Box 62, Oak Ridge, TN 37831. Information Services & Use, December 2013, DOI-10.3233/ISU-130711

This paper takes a broad view of the current status of the use and role of Social Networking by scientists and researchers. Facets of the study include how the styles of scientific collaboration differ from current social media capabilities, a look at the present scientific collaboration sites, what the trends are in the marketplace, and how this paradigm has the potential to change the landscape of the traditional refereed journals. The study is based on a combination of research, technology reviews and statistical analysis of research collaboration using the U. S. Department of Energy-funded published research results.

- *PeerJ – A Case Study in Improving Research Collaboration at the Journal Level*, Peter Binfield, Publisher of *PeerJ*. Information Services & Use, December 2013, DOI-10.3233/ISU-130714

PeerJ was launched with many innovative features which will ultimately serve to improve research collaboration. Owing to the short time the site has been operational, we do not have a large enough database of interactions or published articles to be able to see the full benefit of much of this new thinking. Over the coming months and years it is expected to [?] become clear that some of these innovations have indeed contributed to a new environment which has facilitated and improved research collaboration at all levels.

- *Collaboration When Working with Big Data: Recording and Sharing Analytical Knowledge within and Across Enterprise Data Sources*, Sukumar, Sreenivas R. and Ferrell, Regina K., Oak Ridge National Laboratory, Oak Ridge, TN. Information Services & Use, December 2013, DOI-10.3233/ISU-130712

This paper addresses a specific challenge in the practice of enterprise knowledge management while extracting actionable nuggets from diverse data sources of seemingly related information. In particular, we address the challenge of archiving knowledge gained through collaboration, dissemination, and visualization as part of the data analysis, inference, and decision-making lifecycle.

- *Research Collaboration Tools for the U.S. Department of Defense*, Schwalb, Sandy I., Public Affairs Officer, Defense Technical Information Center , 8725 John J. Kingman Rd., Ft. Belvoir, VA 22630. Information Services & Use, December 2013, DOI-10.3233/ISU-130710

This article examines the approach taken by the Defense Technical Information Center (DTIC) to fulfill a key role by producing secure collaborative tools that facilitate a rapid response to current and emerging threats.

- *Social Network Analysis of Scientific Collaborations across Different Subject Fields*, Bozdogan, Hamparsum and Akbilgic, Oguz, Department of Statistics, Operations, and Management Science, The University of Tennessee, Knoxville, TN 37996. Information Services & Use, December 2013, DOI-10.3233/ISU-130715

This paper analyzes the level of scientific collaboration and interaction in different subject fields using a novel social network analysis (SNA) on a data set provided by the Department of Energy (DOE) Office of Scientific and Technologic Information (OSTI) in Oak Ridge, Tennessee. This paper not only determines the level of scientific collaboration between different disciplines, but it also analyzes the trends among the subject fields considered. The results in this paper show clear pattern recognition discovery on the social network model among key subject fields using weighted networks.

In addition, the PI on the grant and the CEO of Information International Associates, Inc. were invited to serve as guest editors for the special issue of the journal. In this role, their overview editorial as cited below was also published in the issue.

Overview Editorial, Carroll, Bonnie and Coppock, Edrick G. Information International Associates, Inc., 104 Union Valley Road, Oak Ridge, TN 37831-4219. Information Services & Use, December 2013, DOI-10.3233/ISU-130713.

Papers 1 and 5 above acknowledged the financial support of the Department of Energy in the research described in the articles. In addition, the Informed Librarian Online (www.informedlibrarian.com), selected article 1 as a Featured Article in their January 2014 issue, for their Premium Members to read.

PERSONNEL WHO WORKED ON THE GRANT				
Name	Affiliation	Role on project	Duration of support	Percentage
Hamparsum Bozdogan, PhD	Department of Statistics, Operations, and Management Science, The University of Tennessee, Knoxville, TN	PI for subcontractor	Entire Period of Performance	25%
Oguz Akbilgic, PhD	Department of Statistics, Operations, and Management Science, The University of Tennessee, Knoxville, TN	Post Grad	9/2012 through 7/2013	60%
Bill Thompson	Department of Statistics, Operations, and Management Science, The University of Tennessee, Knoxville, TN	Graduate Student	9/2011 through 6/2012	<10%

Edrick G. Coppock	Information International Associates, Inc., Oak Ridge, TN	Prime PI	Entire Period of Performance	25%
Josh Nelson/Matthew Wilson	Information International Associates, Inc., Oak Ridge, TN	Technical	7/2010 through 12/2011	5%
David Resseguie	Oak Ridge National Laboratory, Oak Ridge, TN	Uncompensated Collaborator	Entire Period of Performance	<5%
Bryan Gorman	Oak Ridge National Laboratory, Oak Ridge, TN	Uncompensated Collaborator	7/2010 through 1/2012	<5%
Mallikarjun Shankar, PhD	Oak Ridge National Laboratory, Oak Ridge, TN	Uncompensated Collaborator	Entire Period of Performance	1%
Debbie O'Dell	Consultant	Subject Expert	7/2010 through 12/2012	20%

Current and Pending Support

Person	Support Update
Edrick Coppock	No other current or pending support
Hamparsum Bozdogan, Ph.D.	No other current or pending support
Mallikarjun Shankar, Ph.D.	No changes
David Resseguie	Current support on ORNL SensorNet Project; no overlap with this project

At present, there is no additional support (current and pending, federal and non-federal) that is related to this research.

Cost status

Costs billed to date (30 March 2014):

Total Billed:

553,707.72

The Final Report and DOE Form 241.3 were submitted to DOE Energy Link System (E-Link) on April 7, 2014.